GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 19:09:51 ; Search time 11133 Seconds
(without alignments)
11192.028 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccgggga.....aataaagttgttgcattccc 2192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2192	100.0	2192	6	CS097362	CS097362 Sequence
	2	2192	100.0	2192	6	E12458	E12458 cDNA encodi
	3	2192	100.0	2192	6	E12459	E12459 cDNA encodi
	4	2192	100.0	2192	6	E12605	E12605 cDNA encodi
	5	2192	100.0	2192	8	AB009010	AB009010 Homo sapi
	6	2191	99.9	2222	8	BC039193	BC039193 Homo sapi
	7	2186.2	99.7	2191	8	AK098827	AK098827 Homo sapi
	8	2181.4	99.5	2191	8	AK129749	AK129749 Homo sapi
	9	2181.4	99.5	2207	8	BC093445	BC093445 Homo sapi
	10	2178.4	99.4	2210	6	CQ979236	CQ979236 Sequence
	11	2177	99.3	2201	6	CQ979237	CQ979237 Sequence
	12	2177	99.3	2201	8	BC000449	BC000449 Homo sapi
	13	2165	98.8	2187	8	BC080583	BC080583 Homo sapi
c	14	2120.4	96.7	121984	8	AC126309	AC126309 Homo sapi
	15	2051.6	93.6	2058	8	AB089613	AB089613 Homo sapi
	16	2031.2	92.7	3996	8	AB209782	AB209782 Homo sapi
	17	1998.8	91.2	2309	6	AX337205	AX337205 Sequence
	18	1998.8	91.2	2309	8	HUMUBI13	M26880 Human ubiqu
	19	1924.8	87.8	2286	8	AB089614	AB089614 Pan trogl
	20	1878.8	85.7	2286	8	AB089616	AB089616 Pongo pyg
	21	1831	83.5	6605	8	HUMPOLYUBI	D63791 Human UbC g
c	22	1801.6	82.2	214228	14	AC153180	AC153180 Bos tauru
	23	1793	81.8	2291	8	AB003730	AB003730 Homo sapi
	24	1747.8	79.7	2817	9	AB003731	AB003731 Cricetulu
	25	1747.8	79.7	3273	9	AB003732	AB003732 Cricetulu
	26	1728.6	78.9	3756	9	AF285161	AF285161 Mus muscu
c	27	1725.4	78.7	177047	14	AC138613	AC138613 Mus muscu
	28	1719	78.4	4212	9	AF285162	AF285162 Mus muscu
	29	1697.2	77.4	1830	8	AB089615	AB089615 Gorilla g
	30	1691.2	77.2	2545	9	RATPILP	D17296 Rattus norv
	31	1687.8	77.0	283826	14	AC113658	AC113658 Rattus no
	32	1607.2	73.3	8730	9	CRUPUQ	D63782 Cricetulus
	33	1600.2	73.0	2120	9	BC021837	BC021837 Mus muscu
c	34	1599.2	73.0	149454	9	AC120412	AC120412 Mus muscu
	35	1597.8	72.9	1728	8	BC008955	BC008955 Homo sapi
	36	1557	71.0	1864	8	AK097415	AK097415 Homo sapi
	37	1556.6	71.0	2083	9	BC089218	BC089218 Rattus no
c	38	1484	67.7	164201	2	AC010047	AC010047 Drosophil
c	39	1484	67.7	309958	2	AE003479	AE003479 Drosophil
	40	1480.8	67.6	3448	2	AY118702	AY118702 Drosophil
c	41	1470.2	67.1	110000	14	CT005272_13	Continuation (14 o
c	42	1466.8	66.9	5309	6	CQ600941	CQ600941 Sequence
c	43	1466.8	66.9	5774	6	CQ600944	CQ600944 Sequence
c	44	1466.8	66.9	5795	6	CQ596561	CQ596561 Sequence
	45	1460.8	66.6	2072	5	BC069831	BC069831 Danio rer

ALIGNMENTS

RESULT 1
CS097362
LOCUS CS097362 2192 bp DNA linear PAT 03-JUN-2005
DEFINITION Sequence 5 from Patent WO2005045070.
ACCESSION CS097362
VERSION CS097362.1 GI:66953811
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Lacroix,B., Krause,A., Puisieux,A. and Bachelot,T.
TITLE Method for prognosticating a breast cancer
JOURNAL Patent: WO 2005045070-A 5 19-MAY-2005;
BIOMERIEUX (FR); Centre Leon Berard (FR)
FEATURES Location/Qualifiers
source 1. .2192
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
```

```
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
```

```
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
RESULT 2
E12458
LOCUS       E12458                  2192 bp    DNA     linear   PAT 27-APR-1998
DEFINITION  cDNA encoding human polyubiquitin.
ACCESSION   E12458
VERSION     E12458.1  GI:3251291
KEYWORDS    JP 1997000262-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2192)
  AUTHORS   Kato,M., Yamaguchi,T., Kin,N. and Sekine,S.
  TITLE     HUMAN UBIQUITIN-BOUND ENZYME E2 AND CDNA CODING THE SAME
  JOURNAL   Patent: JP 1997000262-A 2 07-JAN-1997;
            SAGAMI CHEM RES CENTER
COMMENT     OS   Homo sapiens
            PN   JP 1997000262-A/2
            PD   07-JAN-1997
            PF   21-JUN-1995 JP 1995154615
            PI   KATO MASASHI, YAMAGUCHI TOMOKO, KIN NANJIYUN, SEKINE SHINGO PC
              C12N15/09,C07H21/04,C07K14/46,C12N9/64,C12P21/02,C12Q1/68, PC
            (C12N15/09,
            PC   C12R1:91),(C12N9/64,C12R1:19),(C12P21/02,C12R1:19); CC
            strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            CC   Feature is identified by experimental;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2192
            FT                   /organism='Homo sapiens'
            FT                   /cell_type='lymphoma'
            FT                   /cell_line='U937'
            FT                   /clone='HP00184'
            FT   CDS             69. .2126
            FT                   /product='human polyubiquitin'.
FEATURES              Location/Qualifiers
     source           1. .2192
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2192;  DB 6;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
RESULT 3
E12459
LOCUS       E12459                  2192 bp    DNA     linear   PAT 27-APR-1998
DEFINITION  cDNA encoding human polyubiquitin.
ACCESSION   E12459
VERSION     E12459.1  GI:3251292
KEYWORDS    JP 1997000263-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2192)
  AUTHORS   Kato,M., Saeki,M., Kato,T., Kin,N. and Sekine,S.
  TITLE     POLYUBIQUITIN CDNA AND PRODUCTION OF LABELED UBIQUITIN
  JOURNAL   Patent: JP 1997000263-A 1 07-JAN-1997;
            SAGAMI CHEM RES CENTER
COMMENT     OS   Homo sapiens
            PN   JP 1997000263-A/1
            PD   07-JAN-1997
            PF   21-JUN-1995 JP 1995154616
            PI   KATO MASASHI, SAEKI MIHORO, KATO TAKAE, KIN NANJIYUN, PI
            SEKINE SHINGO
            PC   C12N15/09,C07H21/04,C07K19/00,C12P21/02,C12Q1/68,(C12N15/09,
            PC   C12R1:91),
            PC   (C12P21/02,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2192
            FT                   /organism='Homo sapiens'
            FT                   /cell_type='lymphoma'
            FT                   /cell_line='U937'
            FT                   /clone='HP00184'
            FT   5'UTR           1. .68
            FT   CDS             69. .2126
            FT                   /product='human polyubiquitin' FT   3'UTR
                 2127. .2192.
FEATURES             Location/Qualifiers
     source          1. .2192
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 2192;  DB 6;  Length 2192;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
```

```
Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 4
E12605

```
LOCUS       E12605                  2192 bp    DNA     linear   PAT 27-APR-1998
DEFINITION  cDNA encoding human polyubiquitin.
ACCESSION   E12605
VERSION     E12605.1  GI:3251437
KEYWORDS    JP 1997037779-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2192)
  AUTHORS   Kato,M., Saeki,M. and Sekine,S.
  TITLE     UBIQUITIN BOND ENZYME E2 AND HUMAN CDNA CODING FOR THE SAME
  JOURNAL   Patent: JP 1997037779-A 2 10-FEB-1997;
            SAGAMI CHEM RES CENTER
COMMENT     OS   Homo sapiens (human)
            PN   JP 1997037779-A/2
            PD   10-FEB-1997
            PF   31-JUL-1995 JP 1995194376
            PI   KATO MASASHI, SAEKI MIHORO, SEKINE SHINGO
            PC   C12N9/00,C07H21/04,C07K14/46,C07K19/00,C12N15/09//C12N1/21, PC
              (C12N9/00,
            PC   C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2192
            FT                   /organism='Homo sapiens'
            FT                   /tissue_type='Lymphoma'
            FT                   /cell_line='U937'
            FT                   /clone='HP00184'
            FT   5'UTR           1. .68
            FT   CDS             69. .2126
            FT                   /product='Human polyubiquitin' FT   3'UTR
            2127. .2192.
FEATURES            Location/Qualifiers
     source         1. .2192
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2192;  DB 6;  Length 2192;
```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
```

```
Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 5
AB009010
LOCUS       AB009010               2192 bp    mRNA    linear   PRI 26-JUL-2001
DEFINITION  Homo sapiens mRNA for polyubiquitin UbC, complete cds.
ACCESSION   AB009010
VERSION     AB009010.1  GI:2647407
KEYWORDS    polyubiquitin UbC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,N., Yamaguchi,T., Sekine,S., Saeki,M., Iwamuro,S. and Kato,S.
  TITLE     Cloning of human polyubiquitin cDNAs and a ubiquitin-binding assay
            involving its in vitro translation product
  JOURNAL   J. Biochem. 125, 35-39 (1998)
REFERENCE   2  (bases 1 to 2192)
  AUTHORS   Kato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-1997) Seishi Kato, Research Institute of National
            Rehabilitation Center for the Disabled, Department of
            Rehabilitation Engineering; 4-1 Namiki, Tokorozawa, Saitama
            359-8555, Japan (E-mail:seishi@rehab.go.jp,
            Tel:042-995-3100(ex.2568), Fax:042-995-3132)
FEATURES             Location/Qualifiers
     source          1. .2192
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HP00184"
     gene            join(31. .48,347,1022,1325,1898)
                     /gene="UbC2"
     variation       31
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="a"
     variation       48
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="t"
     gene            69. .2126
                     /gene="UbC1"
     CDS             69. .2126
                     /gene="UbC1"
                     /codon_start=1
                     /product="polyubiquitin UbC"
                     /protein_id="BAA23632.1"
                     /db_xref="GI:2647408"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGV"
     variation       347
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="c"
     variation       1022
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="t"
     variation       1325
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="c"
     variation       1898
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="a"
     polyA_signal    2173. .2178
ORIGIN

  Query Match             100.0%;  Score 2192;  DB 8;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATGCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATGTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
```

```
Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CGCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 6
BC039193

```
LOCUS       BC039193                2222 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens ubiquitin C, mRNA (cDNA clone MGC:14624
            IMAGE:4076286), complete cds.
ACCESSION   BC039193
VERSION     BC039193.1  GI:24657521
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2222)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2222)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
```

```
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 21 Row: g Column: 17.
FEATURES              Location/Qualifiers
     source           1. .2222
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:14624 IMAGE:4076286"
                      /tissue_type="Brain, primitive neuroectodermal"
                      /clone_lib="NIH_MGC_56"
                      /lab_host="DH10B"
                      /note="Vector: pDNR-LIB"
     gene             1. .2222
                      /gene="UBC"
                      /note="synonym: HMG20"
                      /db_xref="GeneID:7316"
                      /db_xref="MIM:191340"
     CDS              69. .2126
                      /gene="UBC"
                      /codon_start=1
                      /product="ubiquitin C"
                      /protein_id="AAH39193.1"
                      /db_xref="GI:24657522"
                      /db_xref="GeneID:7316"
                      /db_xref="MIM:191340"
                      /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                      FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                      KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                      TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                      LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                      QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                      QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                      ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                      LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                      GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                      GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                      VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                      GGV"
ORIGIN

  Query Match             99.9%;  Score 2191;  DB 8;  Length 2222;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61

Qy         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121

Qy        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181

Qy        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241

Qy        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301

Qy        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361

Qy        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421

Qy        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481

Qy        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541

Qy        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601

Qy        602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661

Qy        662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721

Qy        722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781

Qy        782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841

Qy        842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901

Qy        902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961

Qy        962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021

Qy       1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081

Qy       1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141

Qy       1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201

Qy       1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261

Qy       1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321

Qy       1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381

Qy       1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441

Qy       1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501

Qy       1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561

Qy       1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621

Qy       1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681

Qy       1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741

Qy       1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801

Qy       1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861

Qy       1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921

Qy       1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981

Qy       1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041

Qy       2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101

Qy       2102 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161

Qy       2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
                ||||||||||||||||||||||||||||||
Db       2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 7
AK098827
LOCUS       AK098827               2191 bp    mRNA    linear   PRI 12-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ25961 fis, clone TST09043, highly similar to
            Polyubiquitin UbC.
ACCESSION   AK098827
VERSION     AK098827.1  GI:21758959
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T.,
            Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R.,
            Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H.,
            Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E.,
            Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S.,
            Terashima,Y., Watanabe,M., Suzuki,Y., Hata,H., Nakagawa,K.,
            Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B.,
            Nagai,K., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2191)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
            full insert sequencing: RAB and Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .2191
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TST09043"
                     /tissue_type="testis"
                     /clone_lib="TST"
                     /note="cloning vector: pME18SFL3"
ORIGIN

  Query Match             99.7%;  Score 2186.2;  DB 8;  Length 2191;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2188;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db          1 AGTTCCGTCGCAGCCGGGAGTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 60

Qy         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 120

Qy        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 180

Qy        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 240

Qy        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 300

Qy        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 360

Qy        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 420

Qy        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 480

Qy        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 540

Qy        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 600
```

```
Qy       602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 660

Qy       662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 720

Qy       722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 780

Qy       782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 840

Qy       842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 900

Qy       902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 960

Qy       962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1020

Qy      1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1080

Qy      1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1140

Qy      1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      1141 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAG 1200

Qy      1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1260

Qy      1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1320

Qy      1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1380

Qy      1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1440

Qy      1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      1441 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACAC 1500

Qy      1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1560

Qy      1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1620

Qy      1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1680

Qy      1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1740

Qy      1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1800

Qy      1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1860

Qy      1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1920

Qy      1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1980

Qy      1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2040

Qy      2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2100

Qy      2102 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2160

Qy      2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
            ||||||||||||||||||||||||||||||
Db      2161 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2191
```

RESULT 8
AK129749

```
LOCUS       AK129749               2191 bp    mRNA    linear   PRI 10-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ26238 fis, clone CBR05440, highly similar to
            Rattus norvegicus ubiquitin C (Ubc).
ACCESSION   AK129749
VERSION     AK129749.1  GI:34526359
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
            Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
            Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,
            Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
            Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
            Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2191)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
```

full insert sequencing: RAB and Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .2191
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBR05440"
/tissue_type="brain"
/clone_lib="CBR"
/note="cloning vector: pME18SFL3"
ORIGIN

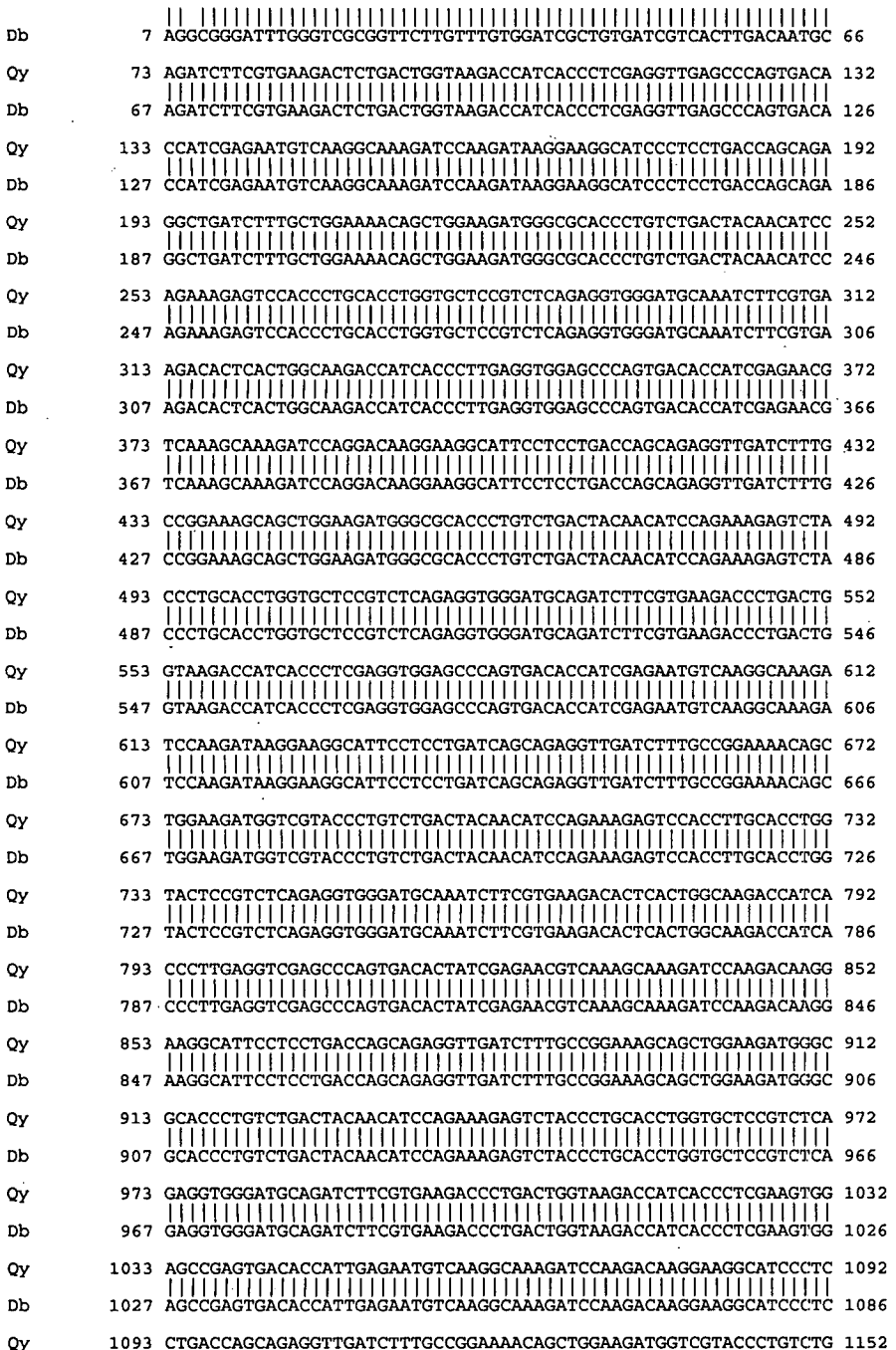
Query Match 99.5%; Score 2181.4; DB 8; Length 2191;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 AGTTCCGTCGCAGCCGGGATTTGGGTCGCAGTTCTTGTTTGTGGATCGCTGTGATCGTCA 60

Qy         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 120

Qy        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 180

Qy        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 240

Qy        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 300

Qy        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        301 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACAC 360

Qy        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 420

Qy        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 480

Qy        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 540

Qy        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 600

Qy        602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 660

Qy        662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 720

Qy        722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 780

Qy        782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 840

Qy        842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 900

Qy        902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 960

Qy        962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1020

Qy       1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1080

Qy       1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1140

Qy       1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1200

Qy       1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1260

Qy       1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1320

Qy       1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1380

Qy       1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1440

Qy       1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1500

Qy       1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1560

Qy       1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1620

Qy       1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1680

Qy       1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1740

Qy       1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CAAGGCAAAGATTCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1800

Qy       1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1860

Qy       1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1861 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGG 1920

Qy       1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1921 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1980

Qy       1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2040

Qy       2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2100

Qy       2102 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2160

Qy       2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||
Db       2161 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2191


RESULT 9
BC093445
LOCUS       BC093445               2207 bp    mRNA    linear   PRI 13-APR-2005
DEFINITION  Homo sapiens cDNA clone IMAGE:5284302, partial cds.
ACCESSION   BC093445
VERSION     BC093445.1  GI:62204947
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2207)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2207)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 168 Row: a Column: 20
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 34304116.
FEATURES             Location/Qualifiers
     source          1. .2207
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5284302"
                     /tissue_type="Brain, hippocampus"
                     /clone_lib="NIH_MGC_95"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     CDS             <1. .2125
                     /codon_start=2
                     /product="Unknown (protein for IMAGE:5284302)"
                     /protein_id="AAH93445.1"
                     /db_xref="GI:62204948"
                     /translation="VPSQPGFGSQFLFVDCCDRHLTMQIFVKTLTGKTITLEVEPSDT
                     IENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIF
                     VKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQ
                     KESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGV"
ORIGIN

  Query Match             99.5%;  Score 2181.4;  DB 8;  Length 2207;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 2185;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61
              |||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||
Db          1 AGTTCCGTCGCAGCCGGGATTTGGGTCGCAGTTCTTGTTTGTGGATTGCTGTGATCGTCA 60

Qy         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 120

Qy        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 180

Qy        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 240

Qy        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 300

Qy        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        301 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACAC 360

Qy        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 420

Qy        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 480
```

```
Qy       482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541
Db       481 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 540

Qy       542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601
Db       541 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 600

Qy       602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661
Db       601 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 660

Qy       662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721
Db       661 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 720

Qy       722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781
Db       721 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 780

Qy       782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841
Db       781 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 840

Qy       842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901
Db       841 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 900

Qy       902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961
Db       901 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 960

Qy       962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021
Db       961 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1020

Qy      1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081
Db      1021 TCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1080

Qy      1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141
Db      1081 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1140

Qy      1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201
Db      1141 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1200

Qy      1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261
Db      1201 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1260

Qy      1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321
Db      1261 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1320

Qy      1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381
Db      1321 TGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1380

Qy      1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441
Db      1381 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1440

Qy      1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501
Db      1441 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1500

Qy      1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561
Db      1501 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1560

Qy      1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621
Db      1561 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1620

Qy      1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681
Db      1621 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1680

Qy      1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741
Db      1681 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1740

Qy      1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801
Db      1741 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1800

Qy      1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861
Db      1801 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1860

Qy      1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921
Db      1861 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGG 1920

Qy      1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981
Db      1921 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1980

Qy      1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041
Db      1981 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2040

Qy      2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101
Db      2041 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2100

Qy      2102 CCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161
Db      2101 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2160

Qy      2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db      2161 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2191
```

```
RESULT 10
CQ979236
LOCUS       CQ979236               2210 bp    DNA     linear   PAT 19-JAN-2005
DEFINITION  Sequence 2 from Patent WO2005001135.
ACCESSION   CQ979236
VERSION     CQ979236.1  GI:57976489
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elander,M.G., Ma,X.J. and Baer,T.M.
  TITLE     3' biased detection of nucleic acids
  JOURNAL   Patent: WO 2005001135-A 2 06-JAN-2005;
            Arcturus Bioscience, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2210
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.4%;  Score 2178.4;  DB 6;  Length 2210;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         13 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 72
```


Db 7 AGGCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 66
Qy 73 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 132
Db 67 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 126
Qy 133 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 192
Db 127 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 186
Qy 193 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 252
Db 187 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 246
Qy 253 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 312
Db 247 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 306
Qy 313 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 372
Db 307 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 366
Qy 373 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
Db 367 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 426
Qy 433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
Db 427 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 486
Qy 493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
Db 487 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 546
Qy 553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
Db 547 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 606
Qy 613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
Db 607 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 666
Qy 673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
Db 667 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 726
Qy 733 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 792
Db 727 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 786
Qy 793 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 852
Db 787 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 846
Qy 853 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 912
Db 847 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 906
Qy 913 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCA 972
Db 907 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCA 966
Qy 973 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGG 1032
Db 967 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGG 1026
Qy 1033 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1092
Db 1027 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1086
Qy 1093 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTG 1152
Db 1087 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTG 1146
Qy 1153 ACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1212
Db 1147 ACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1206
Qy 1213 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACA 1272
Db 1207 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACA 1266
Qy 1273 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGA 1332
Db 1267 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGA 1326
Qy 1333 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1392
Db 1327 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1386
Qy 1393 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1452
Db 1387 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1446
Qy 1453 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1512
Db 1447 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1506
Qy 1513 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1572
Db 1507 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1566
Qy 1573 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1632
Db 1567 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1626
Qy 1633 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1692
Db 1627 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1686
Qy 1693 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1752
Db 1687 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1746
Qy 1753 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1812
Db 1747 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1806
Qy 1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
Db 1807 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1866
Qy 1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
Db 1867 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1926
Qy 1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
Db 1927 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1986
Qy 1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2052
Db 1987 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2046
Qy 2053 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
Db 2047 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2106
Qy 2113 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2172
Db 2107 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2166
Qy 2173 AATAAAGTTGTTGCATTCCC 2192
Db 2167 AATAAAGTTGTTGCATTCCC 2186

```
RESULT 11
CQ979237
LOCUS       CQ979237                2201 bp    DNA     linear   PAT 19-JAN-2005
DEFINITION  Sequence 3 from Patent WO2005001135.
ACCESSION   CQ979237
VERSION     CQ979237.1  GI:57976490
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elander,M.G., Ma,X.J. and Baer,T.M.
  TITLE     3' biased detection of nucleic acids
  JOURNAL   Patent: WO 2005001135-A 3 06-JAN-2005;
            Arcturus Bioscience, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2201
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.3%;  Score 2177;  DB 6;  Length 2201;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740
```

```
Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
             |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 12
BC000449

```
LOCUS       BC000449                2201 bp    mRNA    linear   PRI 06-NOV-2003
DEFINITION  Homo sapiens ubiquitin C, mRNA (cDNA clone IMAGE:2821375), partial
            cds.
ACCESSION   BC000449
VERSION     BC000449.2  GI:38197156
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2201)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2201)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2000) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Nov 6, 2003 this sequence version replaced gi:12653358.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 1 Row: f Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 20149305.
FEATURES             Location/Qualifiers
     source          1. .2201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2821375"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            <1. .2201
                     /gene="UBC"
                     /note="synonym: HMG20"
                     /db_xref="GeneID:7316"
                     /db_xref="MIM:191340"
     CDS             <1. .2111
                     /gene="UBC"
                     /codon_start=3
                     /product="UBC protein"
                     /protein_id="AAH00449.2"
                     /db_xref="GI:38197157"
                     /db_xref="GeneID:7316"
                     /db_xref="MIM:191340"
                     /translation="GFGSRFLFVDRCDRHLTMQIFVKTLTGKTITLEVEPSDTIENVK
                     AKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLT
                     GKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTL
                     HLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ
                     LEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQ
                     DKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTI
                     TLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVL
                     RLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDG
                     RTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG
                     IPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEV
                     EPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
                     GMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLS
                     DYNIQKESTLHLVLRLRGGV"
     misc_feature    54. .275
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    282. .503
                     /gene="UBC"
```

```
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    510. .731
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    738. .959
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    966. .1187
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    1194. .1415
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    1422. .1643
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    1650. .1871
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    1878. .2099
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
ORIGIN

  Query Match           99.3%;  Score 2177;  DB 8;  Length 2201;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
             |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

```
RESULT 13
BC080583
LOCUS       BC080583               2187 bp    mRNA    linear   PRI 07-SEP-2004
DEFINITION  Homo sapiens cDNA clone IMAGE:2822684, partial cds.
ACCESSION   BC080583
VERSION     BC080583.1  GI:51873889
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2187)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  PUBMED    12477932
REFERENCE   2  (bases 1 to 2187)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 59 Row: a Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 34304116.
FEATURES             Location/Qualifiers
     source          1. .2187
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2822684"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             <1. .2099
                     /codon_start=3
                     /product="Unknown (protein for IMAGE:2822684)"
                     /protein_id="AAH80583.1"
                     /db_xref="GI:51873890"
                     /translation="RFLFVDRCDRHLTMQIFVKTLTGKTITLEVEPSDTIENVKAKIQ
                     DKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTI
                     TLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVL
                     RLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDG
                     RTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG
                     IPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEV
                     EPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
                     GMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLS
                     DYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPD
```

QQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSD
TIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQI
FVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNI
QKESTLHLVLRLRGGV"

ORIGIN

Query Match 98.8%; Score 2165; DB 8; Length 2187;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy      88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy     148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy     208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy     268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy     328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy     388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy     448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy     508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy     568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy     628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy     688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 720

Qy     748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 780

Qy     808 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 840

Qy     868 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 900

Qy     928 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 960

Qy     988 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 1020

Qy    1048 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1080

Qy    1108 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 1167
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 1140

Qy    1168 AAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 1227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 AAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 1200

Qy    1228 CCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCA 1287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCA 1260

Qy    1288 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTG 1347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTG 1320

Qy    1348 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1407
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1380

Qy    1408 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1467
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1440

Qy    1468 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1527
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1500

Qy    1528 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1587
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1560

Qy    1588 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1647
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1620

Qy    1648 TCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1707
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 TCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1680

Qy    1708 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1767
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1740

Qy    1768 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1827
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1800

Qy    1828 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 1887
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 1860

Qy    1888 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1947
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1920

Qy    1948 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 2007
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 1980

Qy    2008 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2067
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2040
```

```
Qy       2068 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2127
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2100

Qy       2128 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2187
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2160

Qy       2188 TTCCC 2192
             |||||
Db       2161 TTCCC 2165
```

RESULT 14
AC126309/c
LOCUS AC126309 121984 bp DNA linear PRI 04-APR-2003
DEFINITION Homo sapiens 12 BAC RP11-592O2 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC126309
VERSION AC126309.8 GI:22003948
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121984)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 6 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jul 30, 2002 this sequence version replaced gi:21913328.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES Location/Qualifiers
source 1. .121984
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-592O2"

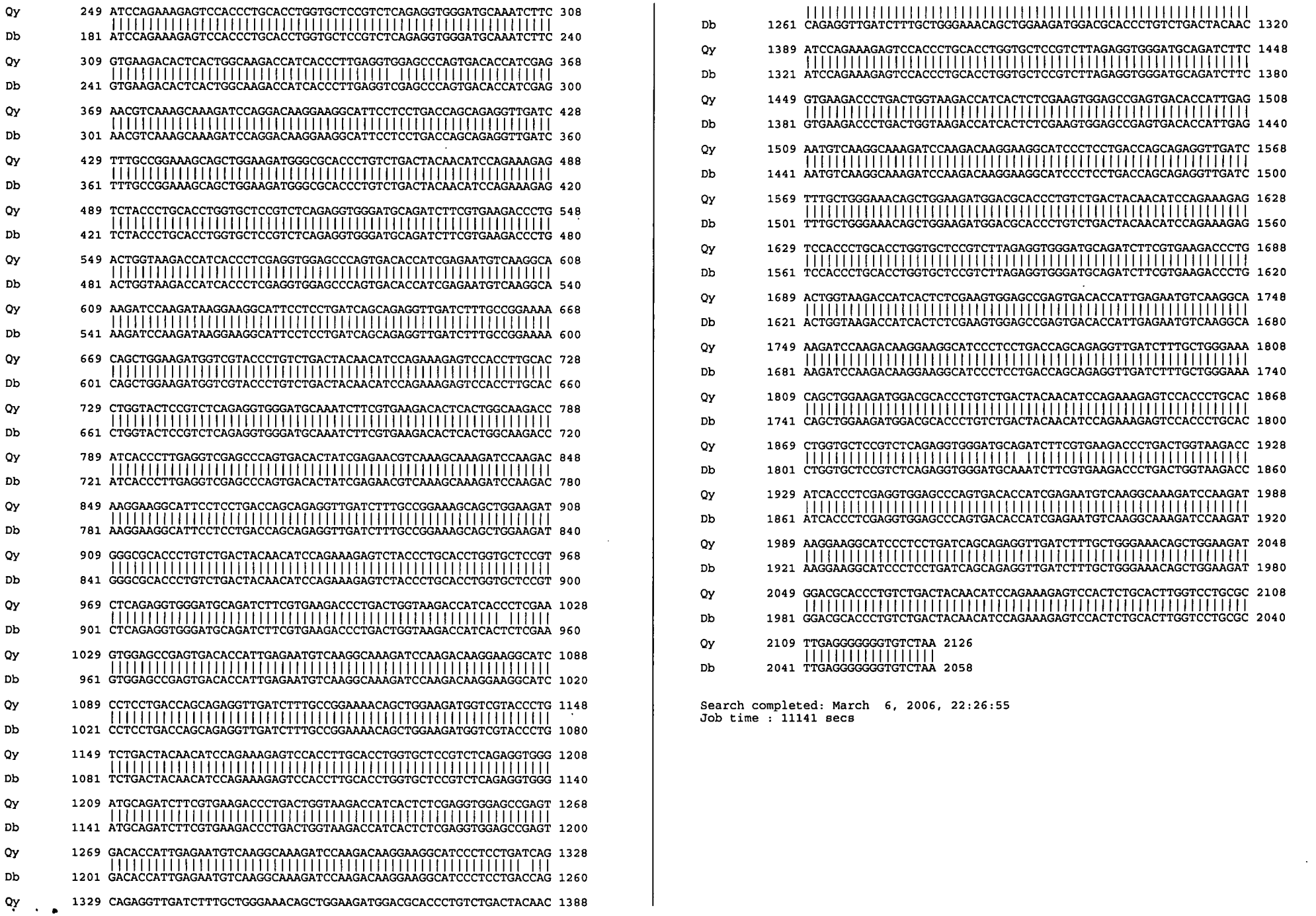
misc_feature	complement(1. .2001)	
	/note="overlaps bases 1. .2001 of clone AC073593"	
	/function="clone overlap"	
repeat_region	319. .510	
	/rpt_family="L2"	
repeat_region	511. .637	
	/rpt_family="FLAM_C"	
repeat_region	638. .1013	
	/rpt_family="L2"	
repeat_region	1191. .1351	
	/rpt_family="Charlie5"	
repeat_region	1353. .1625	
	/rpt_family="L1MC4"	
STS	1583. .1866	
	/standard_name="G59354"	
repeat_region	1641. .1772	
	/rpt_family="FLAM_C"	
STS	2081. .2379	
	/standard_name="SHGC-84469"	
repeat_region	2829. .2869	
	/rpt_family="MIR"	
repeat_region	complement(3610. .3910)	
	/rpt_family="AluSq"	
repeat_region	4545. .4911	
	/rpt_family="MLT1I"	
repeat_region	5210. .5291	
	/rpt_family="MIR"	
repeat_region	5369. .5609	
	/rpt_family="MLT1J"	
repeat_region	5626. .5828	
	/rpt_family="MLT1J"	
repeat_region	6299. .6421	
	/rpt_family="MLT1K"	
repeat_region	7696. .7834	
	/rpt_family="MIR"	
repeat_region	7869. .7920	
	/rpt_family="L2"	
repeat_region	7921. .8217	
	/rpt_family="AluSx"	
repeat_region	8218. .8409	
	/rpt_family="L2"	
repeat_region	8534. .8692	
	/rpt_family="L2"	
repeat_region	8711. .8801	
	/rpt_family="L2"	
repeat_region	complement(9119. .9205)	
	/rpt_family="MIR"	
repeat_region	9358. .9504	
	/rpt_family="(TG)n"	
repeat_region	9585. .9716	
	/rpt_family="FLAM_C"	
repeat_region	complement(10426. .10580)	
	/rpt_family="L2"	
repeat_region	10671. .10797	
	/rpt_family="AluJ"	
repeat_region	complement(10799. .10968)	
	/rpt_family="AluYa8"	
repeat_region	complement(10971. .11291)	
	/rpt_family="AluSx"	
repeat_region	11292. .11493	
	/rpt_family="AluJ"	
repeat_region	12190. .12250	
	/rpt_family="MIR"	
repeat_region	complement(12256. .12380)	
	/rpt_family="MER102"	
STS	12488. .12761	
	/standard_name="SHGC-111902"	
repeat_region	complement(12595. .12666)	
	/rpt_family="MIR"	

Query Match 96.7%; Score 2120.4; DB 8; Length 121984;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2124; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy        63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106114 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 106055

Qy       123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106054 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 105995

Qy       183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105994 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 105935

Qy       243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105934 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 105875

Qy       303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    105874 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACC 105815

Qy       363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105814 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 105755

Qy       423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105754 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 105695

Qy       483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105694 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 105635

Qy       543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105634 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 105575

Qy       603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    105574 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTTCTGATCAGCAGAGGTTGATCTTTGCC 105515

Qy       663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105514 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 105455

Qy       723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105454 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 105395

Qy       783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105394 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 105335

Qy       843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105334 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 105275

Qy       903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105274 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 105215

Qy       963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105214 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 105155

Qy      1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105154 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 105095

Qy      1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105094 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 105035
```

```
Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     105034 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 104975

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104974 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 104915

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104914 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 104855

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104854 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 104795

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104794 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 104735

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104734 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 104675

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104674 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 104615

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104614 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 104555

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104554 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 104495

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104494 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 104435

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104434 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 104375

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104374 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 104315

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
             ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     104314 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGT 104255

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104254 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 104195

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104194 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 104135

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104134 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 104075

Qy       2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     104074 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 104015

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||||
Db     104014 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 103985
```

```
RESULT 15
AB089613
LOCUS       AB089613               2058 bp    DNA     linear   PRI 03-FEB-2004
DEFINITION  Homo sapiens UbC gene for polyubiquitin C, complete cds.
ACCESSION   AB089613
VERSION     AB089613.1  GI:28316335
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tachikui,H., Saitou,N., Nakajima,T., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Lineage-specific homogenization of the polyubiquitin gene among
            human and great apes
  JOURNAL   J. Mol. Evol. 57 (6), 737-744 (2003)
   PUBMED   14745543
REFERENCE   2  (bases 1 to 2058)
  AUTHORS   Tachikui,H., Nakajima,T., Saito,N., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2002) Hiroshi Tachikui, Institute of Medical
            Science, University of tokyo, Division of Genetic Dignosis;
            Shirokanedai 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:tachikui@ims.u-tokyo.ac.jp, Tel:81-5449-5750,
            Fax:81-5449-5764)
FEATURES             Location/Qualifiers
     source          1. .2058
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Japanese with 9 ub units"
     gene            1. .2058
                     /gene="UbC"
     CDS             1. .2058
                     /gene="UbC"
                     /codon_start=1
                     /product="polyubiquitin C"
                     /protein_id="BAC56951.1"
                     /db_xref="GI:28316336"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGV"
ORIGIN

  Query Match           93.6%;  Score 2051.6;  DB 8;  Length 2058;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2054;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 60

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 120

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 180
```


Qy 249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
Db 181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 240
Qy 309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
Db 241 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAG 300
Qy 369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
Db 301 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 360
Qy 429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
Db 361 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 420
Qy 489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
Db 421 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 480
Qy 549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
Db 481 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 540
Qy 609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
Db 541 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 600
Qy 669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
Db 601 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 660
Qy 729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
Db 661 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 720
Qy 789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
Db 721 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 780
Qy 849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
Db 781 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 840
Qy 909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
Db 841 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 900
Qy 969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028
Db 901 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 960
Qy 1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
Db 961 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1020
Qy 1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
Db 1021 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1080
Qy 1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
Db 1081 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1140
Qy 1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
Db 1141 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1200
Qy 1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
Db 1201 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1260
Qy 1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
Db 1261 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1320
Qy 1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448
Db 1321 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1380
Qy 1449 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1508
Db 1381 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1440
Qy 1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
Db 1441 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1500
Qy 1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
Db 1501 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1560
Qy 1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
Db 1561 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1620
Qy 1689 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1748
Db 1621 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1680
Qy 1749 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1808
Db 1681 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1740
Qy 1809 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1868
Db 1741 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1800
Qy 1869 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1928
Db 1801 CTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACC 1860
Qy 1929 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 1988
Db 1861 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 1920
Qy 1989 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 2048
Db 1921 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 1980
Qy 2049 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2108
Db 1981 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2040
Qy 2109 TTGAGGGGGGGTGTCTAA 2126
Db 2041 TTGAGGGGGGGTGTCTAA 2058

Search completed: March 6, 2006, 22:26:55
Job time : 11141 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 18:56:52 ; Search time 1221 Seconds
(without alignments)
11964.788 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2192	100.0	2192	2	AAT62351	Aat62351 Human pol
2	2192	100.0	2192	2	AAT64489	Aat64489 Poly-Ubiq
3	2192	100.0	2192	6	ABK83453	Abk83453 Human cDN
4	2192	100.0	2192	6	ABV78062	Abv78062 Hypoxia-r
5	2192	100.0	2192	10	ADD14753	Add14753 Human src
6	2192	100.0	2192	12	ADP02960	Adp02960 Human hou
7	2192	100.0	2192	13	ADS88458	Ads88458 Human hou
8	2192	100.0	2192	13	ADU60040	Adu60040 Housekeep
9	2192	100.0	2192	14	ADZ67404	Adz67404 Different
10	2178.4	99.4	2210	14	ADW02234	Adw02234 Human ubi
11	2177	99.3	2201	14	ADW02235	Adw02235 Human ubi
12	2128.4	97.1	3055	13	ADQ84521	Adq84521 Human tum
13	2128.4	97.1	3055	13	ACN39601	Acn39601 Tumour-as
14	1998.8	91.2	2309	6	ABL69377	Abl69377 Prostate
15	1998.8	91.2	2309	6	ABK83973	Abk83973 Human cDN
16	1998.8	91.2	2309	14	ADX05493	Adx05493 Cyclin-de
17	1928.4	88.0	2058	14	ADV42626	Adv42626 Human psy
18	1793	81.8	2291	10	ABT17058	Abt17058 Human MP2
19	1691.2	77.2	2545	10	ADB52865	Adb52865 Primary r
20	1691.2	77.2	2545	10	ABT41997	Abt41997 Toxicity
c 21	1466.8	66.9	5309	4	ABL20972	Abl20972 Drosophil
c 22	1466.8	66.9	5774	4	ABL20974	Abl20974 Drosophil
c 23	1466.8	66.9	5795	4	ABL18052	Abl18052 Drosophil
24	1276.8	58.2	7648	4	AAH48763	Aah48763 Vector pH
25	1250.2	57.0	2121	13	ADX32148	Adx32148 Plant ful
26	1230.4	56.1	2419	4	ABL20973	Abl20973 Drosophil
27	1230.4	56.1	2428	4	ABL18053	Abl18053 Drosophil
28	1146.2	52.3	1793	13	ADX60958	Adx60958 Plant ful
29	1144.6	52.2	1893	13	ADX62486	Adx62486 Plant ful
30	1143	52.1	1790	13	ADX59863	Adx59863 Plant ful
31	1143	52.1	1885	13	ADX60408	Adx60408 Plant ful
32	1143	52.1	1896	13	ADX61692	Adx61692 Plant ful
33	1143	52.1	1902	13	ADX62442	Adx62442 Plant ful
34	1143	52.1	2110	13	ADX11132	Adx11132 Plant ful
35	1140.4	52.0	1852	13	ADX45557	Adx45557 Plant ful
36	1138.8	52.0	1813	13	ADX61500	Adx61500 Plant ful
37	1138.8	52.0	1846	13	ADX60054	Adx60054 Plant ful
38	1138.8	52.0	1846	13	ADX60029	Adx60029 Plant ful
39	1138.8	52.0	1849	13	ADX46673	Adx46673 Plant ful
40	1138.8	52.0	1850	13	ADX59544	Adx59544 Plant ful
41	1138.8	52.0	1854	13	ADX62417	Adx62417 Plant ful
42	1138.8	52.0	1854	13	ADX61895	Adx61895 Plant ful
43	1138.8	52.0	1857	13	ADX63630	Adx63630 Plant ful
44	1138.8	52.0	1858	13	ADX35574	Adx35574 Plant ful
45	1138.8	52.0	1858	13	ADX11086	Adx11086 Plant ful

ALIGNMENTS

RESULT 1
AAT62351
ID AAT62351 standard; cDNA to mRNA; 2192 BP.
XX
AC AAT62351;
XX
DT 19-JUN-1997 (first entry)
XX
DE Human poly-ubiquitin cDNA sequence.
XX
KW Human; ubiquitin; lymphoma; in vitro translation; diagnosis; disease;
KW rabbit reticulocyte lysate; metabolism; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 69. .2126
FT /*tag= a
FT /product= "human poly-ubiquitin protein"
FT /note= "CDS is claimed in Claim 1"
XX
PN JP09000263-A.
XX
PD 07-JAN-1997.
XX
PF 21-JUN-1995; 95JP-00154616.
XX
PR 21-JUN-1995; 95JP-00154616.
XX
PA (SAGA) SAGAMI CHEM RES CENTRE.
XX
DR WPI; 1997-112846/11.
DR P-PSDB; AAW14134.
XX
PT Human poly-ubiquitin cDNA - for prodn. of labelled poly-ubiquitin, useful
PT for diagnosing diseases resulting from abnormalities of ubiquitin
PT metabolism.
XX
PS Claim 2; Page 8-10; 11pp; Japanese.
XX
CC This nucleotide sequence encodes the human poly-ubiquitin protein. The
CC gene was isolated from a cDNA library generated from polyA+ purified from

CC the human lymphoma cell line U937. Confirmation that the sequence encoded
CC the correct protein was done by in vitro translation in a rabbit
CC reticulocyte lysate. The DNA and protein can be used for diagnosis of
CC certain diseases resulting from abnormalities in metabolism of ubiquitins
CC or as a reagent for study of the ubiquitin system
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 2; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 2
AAT64489
ID   AAT64489 standard; cDNA to mRNA; 2192 BP.
XX
AC   AAT64489;
XX
DT   03-JUN-1997  (first entry)
XX
DE   Poly-Ubiquitin cDNA.
XX
KW   Human; ubiquitin-combining enzyme E2;
KW   epidermal cancer cell line KB ATCC CRL 17; ds.
XX
OS   Synthetic.
XX
PN   JP09037779-A.
XX
PD   10-FEB-1997.
XX
PF   31-JUL-1995;   95JP-00194376.
XX
PR   31-JUL-1995;   95JP-00194376.
XX
PA   (SAGA ) SAGAMI CHEM RES CENTRE.
XX
DR   WPI; 1997-173039/16.
DR   P-PSDB; AAW14848.
XX
PT   Human ubiquitin-combining enzyme E2 - derived from human epidermal cancer
PT   cell line KB ATCC CRL 17; useful as a research reagent.
XX
PS   Example; Page 7-10; 10pp; Japanese.
XX
CC   This sequence encodes polyubiquitin which was used to demonstrate the
CC   ubiquitin binding properties of the human ubiquitin-combining enzyme E2.
CC   Human ubiquitin-combining enzyme E2 and cDNA encoding it are useful as
CC   tools in biological research. The human E2 protein was isolated from
CC   human epidermal cancer cell line KB ATCC CRL 17
XX
SQ   Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 2192;  DB 2;  Length 2192;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 3
ABK83453
ID ABK83453 standard; cDNA; 2192 BP.
XX
AC ABK83453;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #24.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 24; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              |||||||||||||||||||||||||||||||
```

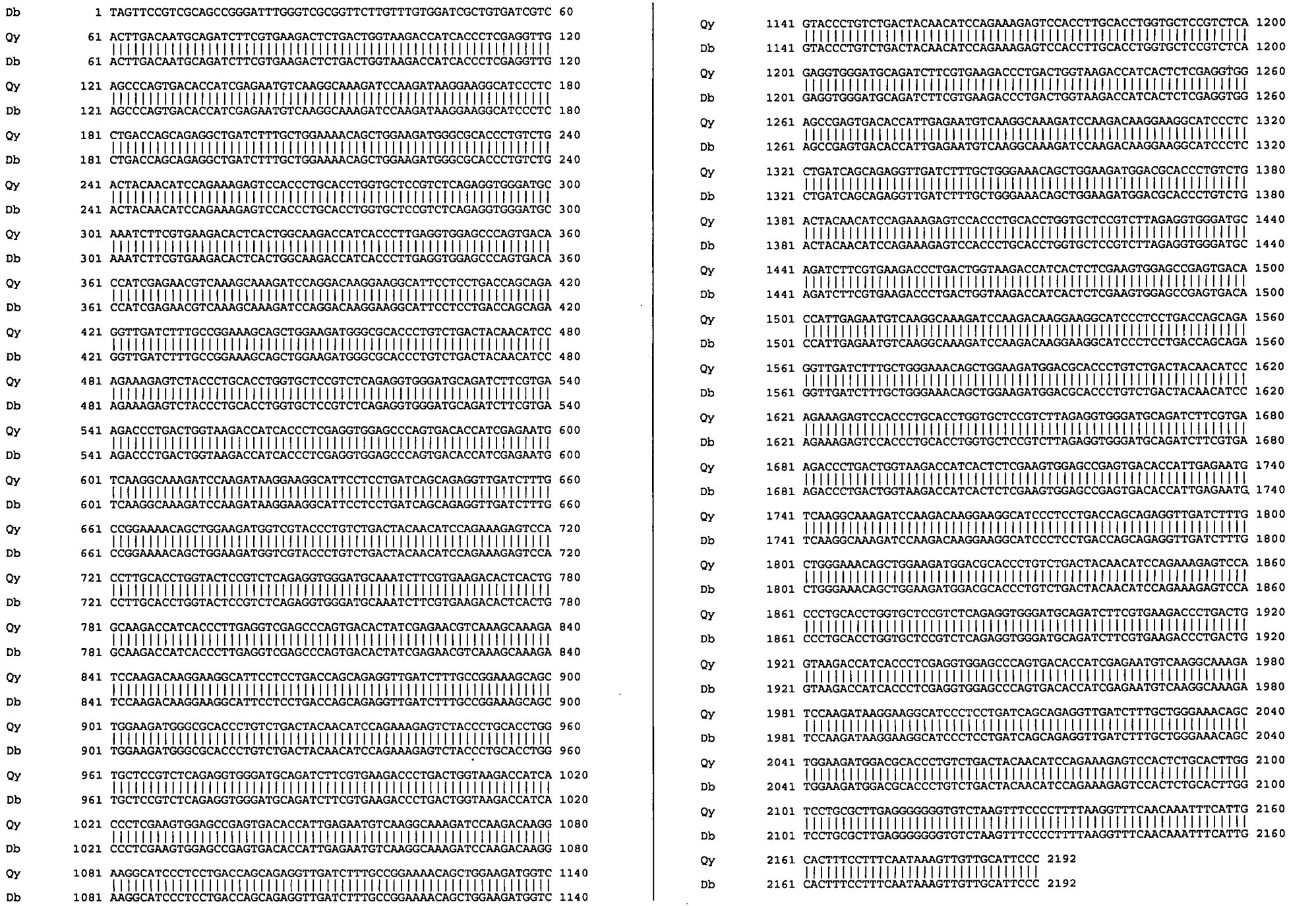
```
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 4
ABV78062
ID   ABV78062 standard; DNA; 2192 BP.
XX
AC   ABV78062;
XX
DT   12-NOV-2002  (first entry)
XX
DE   Hypoxia-regulated protein coding sequence #82.
XX
KW   Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW   antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
KW   hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
KW   ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW   preeclapmsia; atherosclerosis; inflammatory condition; wound healing;
KW   inflammation; erythropoiesis; hair loss; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200246465-A2.
XX
PD   13-JUN-2002.
XX
PF   10-DEC-2001; 2001WO-GB005458.
XX
PR   08-DEC-2000; 2000GB-00030076.
PR   08-FEB-2001; 2001GB-00003156.
PR   25-OCT-2001; 2001GB-00025666.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   White J,  Mundy CR,  Ward NR,  Krige D,  Kingsman SM,  Harris RA;
PI   Rayner WN;
XX
DR   WPI; 2002-627238/67.
XX
PT   Identifying a gene involved in disease for treating hypoxia-regulated
PT   conditions, comprises comparing the transcriptome/proteome of two cell
PT   types under different conditions and identifying a differentially
PT   regulated gene.
XX
PS   Claim 37; Page 393; 538pp; English.
XX
CC   The present invention relates to methods for identifying genes and
CC   proteins that are implicated in a specific disease or physiological
CC   condition. The method comprises comparing the transcriptome/proteome of a
CC   specialised cell type implicated in a disease or condition with that of a
CC   second specialised cell type, under two experimental conditions, and
CC   identifying a gene that is differentially regulated in the two
CC   specialised cell types under experimental conditions. ABV77873-ABV78116
CC   and ABP65061-ABP65257 were identified using the methods of the invention.
CC   The coding sequences and proteins are useful for treating a disease in a
CC   patient, for manufacture of a medicament for treating hypoxia-regulated
CC   conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,
CC   biological response to hypoxia conditions, or hypoxic-associated
CC   pathology in a patient. The coding sequences and proteins are also useful
CC   for monitoring the therapeutic treatment of a disease or physiological
CC   condition, such as cancer, ischaemic conditions, reperfusion injury,
CC   retinopathy, neonatal stress, preeclapmsia, atherosclerosis, inflammatory
CC   conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
SQ   Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 2192;  DB 6;  Length 2192;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
```

```
Qy      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 5
ADD14753
ID ADD14753 standard; cDNA; 2192 BP.
XX
AC ADD14753;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polynucleotide SEQ ID NO:147.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003062395-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001981.
XX
PR 18-JAN-2002; 2002US-0350061P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX
DR WPI; 2003-636735/60.
DR P-PSDB; ADD14151.
XX
PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
PS Claim 2; SEQ ID NO 147; 139pp; English.
XX
CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 10; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
Qy 61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Db 61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Qy 121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Db 121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Qy 181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Db 181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Qy 241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Db 241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Qy 301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db 301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Qy 361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db 361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Qy 421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db 421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Qy 481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db 481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Qy 541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db 541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Qy 601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db 601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Qy 661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db 661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Qy 721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db 721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Qy 781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db 781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Qy 841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db 841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Qy 901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db 901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Qy 961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db 961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Qy 1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db 1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Qy 1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db 1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Qy 1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db 1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Qy 1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db 1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Qy 1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db 1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Qy 1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db 1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Qy 1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db 1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Qy 1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db 1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Qy 1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db 1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Qy 1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db 1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Qy 1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db 1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Qy 1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db 1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Qy 1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db 1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Qy 1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db 1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Qy 1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db 1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Qy 1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db 1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Qy 1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db 1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Qy 2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db 2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Qy 2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Db 2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Qy 2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db 2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192

```
RESULT 6
ADP02960
ID   ADP02960 standard; cDNA; 2192 BP.
XX
AC   ADP02960;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human housekeeping gene cDNA #1.
XX
KW   ss; gene; human; housekeeping gene; capture probe; DNA microarray;
KW   cancer.
XX
OS   Homo sapiens.
XX
PN   JP2004135552-A.
XX
PD   13-MAY-2004.
XX
PF   16-OCT-2002; 2002JP-00302278.
XX
PR   16-OCT-2002; 2002JP-00302278.
XX
PA   (NIGA ) NGK INSULATORS LTD.
XX
DR   WPI; 2004-404613/38.
XX
PT   Novel human housekeeping gene commonly expressed in different human
PT   tissues, useful as capture probe in DNA microarray.
XX
PS   Claim 1; SEQ ID NO 1; 940pp; Japanese.
XX
CC   The invention relates to a human housekeeping gene. The housekeeping gene
CC   is useful as a capture probe in DNA microarray. The housekeeping gene is
CC   useful as a reference with respect to the normal gene expression measured
CC   on various conditions, and an expression database of normal human tissue
CC   is useful as reference information for using a disease related gene for
CC   diagnosis or treatment. The housekeeping gene when used as capture probe,
CC   enables diagnosis of cancer. The present sequence represents a human
CC   housekeeping gene cDNA.
XX
SQ   Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2192;  DB 12;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 7
ADS88458
ID ADS88458 standard; cDNA; 2192 BP.
XX
AC ADS88458;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human housekeeping gene cDNA sequence SEQ ID NO:1.
XX
KW housekeeping gene; human; probe; hybridisation; DNA microarray;
KW disease investigation; disease diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004035785-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2002; 2002WO-JP010753.
XX
PR 16-OCT-2002; 2002WO-JP010753.
XX
PA (NIGA) NGK INSULATORS LTD.
XX
PI Aburatani H, Yamamoto S;
XX
DR WPI; 2004-357219/33.
XX
PT Human housekeeping and tissue-specific genes, gene sets and transcription
PT products, probes and microarrays derived from them for reference use in
PT investigation and diagnosis of disease.
XX
PS Claim 5; SEQ ID NO 1; 372pp; Japanese.
XX
CC The present invention describes 1189 housekeeping genes and gene sets
CC containing them, expressed in 35 different human tissues. Also described:
CC (1) RNA and cDNA transcription products of the housekeeping genes; (2)
CC oligonucleotide probes hybridising to the housekeeping genes; (3) DNA
CC microarrays containing the oligonucleotide probes; (4) 1704 genes
CC expressed specifically in particular tissues; (5) RNA and cDNA
CC transcription products of the tissue-specific genes; (6) oligonucleotide
CC probes hybridising to the tissue-specific genes; and (7) DNA microarrays
CC containing the oligonucleotide probes. The tissue-specific genes are
CC expressed specifically in one of brain, thymus, pituitary, thyroid,
CC trachea, lung, breast, skin, skeletal muscle, heart, liver, spleen,
CC kidney, adrenal gland, pancreas, stomach, small intestine, colon,
CC bladder, prostate gland, testis, ovary, placenta, uterus, bone marrow,
CC foetal brain or foetal liver. The housekeeping genes can be used as
CC reference genes for the investigation and diagnosis of disease. The
CC present sequence represents a human housekeeping gene cDNA sequence which
CC is given in the exemplification of the present invention.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 13; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
```

```
Qy       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
            ||||||||||||||||||||||||||||||||
Db      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 8
ADU60040
ID ADU60040 standard; cDNA; 2192 BP.
XX
AC ADU60040;
XX
DT 10-FEB-2005 (first entry)
XX
DE Housekeeping gene cDNA, SEQ ID 1.
XX
KW Housekeeping; gene; ss.
XX
OS Homo sapiens.
XX
PN US2004229233-A1.
XX
PD 18-NOV-2004.
XX
PF 15-OCT-2003; 2003US-00684422.
XX
PR 16-OCT-2002; 2002US-0418614P.
XX
PA (NIGA) NGK INSULATORS LTD.
XX
PI Aburatani H, Yamamoto S;
XX
DR WPI; 2004-813237/80.
XX
PT New human housekeeping genes and tissue-specific genes, used as a
PT reference for a normal or disease-associated gene expression measured
PT under various conditions.
XX
PS Claim 5; SEQ ID NO 1; 93pp; English.

XX
CC The present invention relates to novel human housekeeping genes and
CC tissue-specific genes. The present sequence is a transcription product of
CC one such gene. Also claimed is a DNA microarray carrying the set of
CC transcription products of the genes of the invention. The housekeeping
CC genes and tissue-specific genes are useful as references for a normal
CC gene expression or disease-associated gene expression measured under
CC various conditions. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC from seqdata.uspto.gov/sequence.html?DocID=20040229233.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 13; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
```

```
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 9
ADZ67404
ID   ADZ67404 standard; cDNA; 2192 BP.
XX
AC   ADZ67404;
XX
DT   14-JUL-2005  (first entry)
XX
DE   Differentially expressed gene #5 in breast cancer relapse risk patient.
XX
KW   prognosis; breast tumor; cancer; gene expression; ss.
XX
OS   Homo sapiens.
XX
PN   FR2861744-A1.
XX
PD   06-MAY-2005.
XX
PF   05-NOV-2003; 2003FR-00012966.
XX
PR   05-NOV-2003; 2003FR-00012966.
XX
PA   (INMR ) BIOMERIEUX SA.
PA   (BERA-) CENT BERARD LEON.
XX
PI   Lacroix B,  Krause A,  Puisieux A,  Bachelot T;
XX
DR   WPI; 2005-348176/36.
XX
PT   Prognosis of breast cancer, useful for identifying patients at high risk
PT   of relapse, based on analysis of differential expression of specific
PT   genes.
XX
PS   Claim 1; SEQ ID NO 5; 154pp; French.
XX
CC   The invention relates to a method for prognosis of breast cancer
CC   comprising extracting biological material (A) from a patient sample;
CC   contacting (A) with a reagent (R) specific for target genes (TG) having
CC   any of 68 sequences reproduced; and determining expression of TG. The
CC   method is used to identify patients at high risk of relapse, particularly
CC   to assist selection of primary treatment. This sequence corresponds to a
CC   differentially regulated gene found in patients at risk of a relapse into
CC   breast cancer.
XX
SQ   Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 2192;  DB 14;  Length 2192;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
```

```
Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 10
ADW02234
ID ADW02234 standard; DNA; 2210 BP.
XX
AC ADW02234;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human ubiquitin C related nucleotide sequence.
XX
KW DNA detection; breast cancer; ubiquitin C; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2005001135-A2.
XX
PD 06-JAN-2005.
XX
PF 02-JUN-2004; 2004WO-US017572.
XX
PR 03-JUN-2003; 2003US-0475812P.
PR 30-JAN-2004; 2004US-00769476.
XX
PA (ARCT-) ARCTURUS BIOSCIENCE INC.
XX
PI Elander MG, Ma X, Baer TM;
XX
DR WPI; 2005-075586/08.
DR GENBANK; BC000449.
XX
PT New microarray, useful for studying functional gene networks and gene
PT expression as the result of a disease state in a cell, tissue, organ, or
PT subject such as cells that are pre-cancerous, cancerous, or invasive such
PT as breast cancer.
XX
PS Example 2; Page 34-35; 45pp; English.
XX
CC The invention relates to a microarray comprising at least 5
CC oligonucleotide probes, each of 150 nucleotides or less in length, and
CC complementary to at least 10 consecutive nucleotides of an mRNA molecule,
CC where the 10 consecutive nucleotides is, in its entirety, less than 360
CC nucleotides from the site of poly(A) addition of the mRNA molecule.
CC Analyzing gene expression in a cell comprises preparing a polynucleotide
CC comprising gene sequences expressed in the cell and hybridizing the
CC polynucleotide to a microarray above. The cell is from an FFPE sample,
CC preferably a breast cancer cell. The cell has been isolated by
CC microdissection of a cell-containing sample. The method may comprise
CC quantitative PCR amplification of a sequence containing at least 10
CC consecutive nucleotides of an mRNA molecule present in the cell, where
CC the 10 consecutive nucleotides is, in its entirety, less than 360
CC nucleotides from the site of poly(A) addition of the mRNA molecule and
CC where the mRNA is optionally a reference sequence. The microarray is
CC useful for conducting quantitative and qualitative analysis of gene
CC expression and for studying gene expression as the result of a disease
CC state in a cell, tissue, organ, or subject, particularly the expression
CC of genes in cells that are not normal, pre-cancerous, cancerous, or
CC invasive such as breast cancer. The microarray is also useful for
CC studying functional gene networks. The methods are useful for detecting
CC cancer cells (e.g. breast cancer cells), facilitating diagnosis of cancer
CC and the severity of a cancer in a subject, facilitating a determination
CC of the prognosis of a subject, and assessing the responsiveness of the
CC subject to therapy. The present sequence represents a human nucleotide
CC sequence that is similar to ubiquitin C, clone MGC:8448 IMAGE:2821375,
CC that can be detected using the method of the invention.
XX
SQ Sequence 2210 BP; 622 A; 571 C; 575 G; 442 T; 0 U; 0 Other;

Query Match 99.4%; Score 2178.4; DB 14; Length 2210;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         13 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 72
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        7 AGGCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 66

Qy       73 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       67 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 126

Qy      133 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 186

Qy      193 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 252
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 246

Qy      253 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      247 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 306

Qy      313 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 366

Qy      373 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      367 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 426

Qy      433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      427 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 486

Qy      493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      487 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 546

Qy      553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      547 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 606

Qy      613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      607 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 666

Qy      673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      667 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 726

Qy      733 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      727 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 786

Qy      793 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      787 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 846

Qy      853 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      847 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 906

Qy      913 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCA 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      907 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCA 966

Qy      973 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGG 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      967 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGG 1026

Qy     1033 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1092
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1027 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1086

Qy     1093 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTG 1152
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1087 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTG 1146

Qy     1153 ACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1212
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1147 ACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1206

Qy     1213 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACA 1272
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1207 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACA 1266

Qy     1273 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGA 1332
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1267 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGA 1326

Qy     1333 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1327 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1386

Qy     1393 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1387 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1446

Qy     1453 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1447 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1506

Qy     1513 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1572
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1507 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1566

Qy     1573 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1567 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1626

Qy     1633 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1692
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1627 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1686

Qy     1693 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1752
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1687 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1746

Qy     1753 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1812
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1747 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1806

Qy     1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1807 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1866

Qy     1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1867 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1926

Qy     1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1927 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1986

Qy     1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1987 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2046

Qy     2053 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2047 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2106

Qy     2113 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2107 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2166

Qy     2173 AATAAAGTTGTTGCATTCCC 2192
           ||||||||||||||||||||
Db     2167 AATAAAGTTGTTGCATTCCC 2186
```

```
RESULT 11
ADW02235
ID   ADW02235 standard; cDNA; 2201 BP.
XX
AC   ADW02235;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Human ubiquitin C cDNA sequence.
XX
KW   DNA detection; breast cancer; ubiquitin C; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2005001135-A2.
XX
PD   06-JAN-2005.
XX
PF   02-JUN-2004; 2004WO-US017572.
XX
PR   03-JUN-2003; 2003US-0475812P.
PR   30-JAN-2004; 2004US-00769476.
XX
PA   (ARCT-) ARCTURUS BIOSCIENCE INC.
XX
PI   Elander MG,  Ma X,  Baer TM;
XX
DR   WPI; 2005-075586/08.
DR   GENBANK; BC000449.
XX
PT   New microarray, useful for studying functional gene networks and gene
PT   expression as the result of a disease state in a cell, tissue, organ, or
PT   subject such as cells that are pre-cancerous, cancerous, or invasive such
PT   as breast cancer.
XX
PS   Example 2; Page 35; 45pp; English.
XX
CC   The invention relates to a microarray comprising at least 5
CC   oligonucleotide probes, each of 150 nucleotides or less in length, and
CC   complementary to at least 10 consecutive nucleotides of an mRNA molecule,
CC   where the 10 consecutive nucleotides is, in its entirety, less than 360
CC   nucleotides from the site of poly(A) addition of the mRNA molecule.
CC   Analyzing gene expression in a cell comprises preparing a polynucleotide
CC   comprising gene sequences expressed in the cell and hybridizing the
CC   polynucleotide to a microarray above. The cell is from an FFPE sample,
CC   preferably a breast cancer cell. The cell has been isolated by
CC   microdissection of a cell-containing sample. The method may comprise
CC   quantitative PCR amplification of a sequence containing at least 10
CC   consecutive nucleotides of an mRNA molecule present in the cell, where
CC   the 10 consecutive nucleotides is, in its entirety, less than 360
CC   nucleotides from the site of poly(A) addition of the mRNA molecule and
CC   where the mRNA is optionally a reference sequence. The microarray is
CC   useful for conducting quantitative and qualitative analysis of gene
CC   expression and for studying gene expression as the result of a disease
CC   state in a cell, tissue, organ, or subject, particularly the expression
CC   of genes in cells that are not normal, pre-cancerous, cancerous, or
CC   invasive such as breast cancer. The microarray is also useful for
CC   studying functional gene networks. The methods are useful for detecting
CC   cancer cells (e.g. breast cancer cells), facilitating diagnosis of cancer
CC   and the severity of a cancer in a subject, facilitating a determination
CC   of the prognosis of a subject, and assessing the responsiveness of the
CC   subject to therapy. The present sequence represents a human ubiquitin C,
CC   clone IMAGE:2821375, that can be detected using the method of the
CC   invention.
XX
SQ   Sequence 2201 BP; 620 A; 569 C; 570 G; 442 T; 0 U; 0 Other;

  Query Match             99.3%;  Score 2177;  DB 14;  Length 2201;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
             |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 12
ADQ84521
ID ADQ84521 standard; cDNA; 3055 BP.
XX
AC ADQ84521;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1335.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 1335; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing

```
CC   a medicament for the therapeutic treatment or diagnostic detection of a
CC   cell proliferative disorder or cancer. The present sequence represents a
CC   human TAT cDNA sequence from the present invention.
XX
SQ   Sequence 3055 BP; 736 A; 744 C; 874 G; 701 T; 0 U; 0 Other;

  Query Match            97.1%;  Score 2128.4;  DB 13;  Length 3055;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 2129;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 933

Qy        123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 993

Qy        183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        994 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 1053

Qy        243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 1113

Qy        303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1114 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 1173

Qy        363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 1233

Qy        423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 1293

Qy        483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 1353

Qy        543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 1413

Qy        603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 1473

Qy        663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1533

Qy        723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1534 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 1593

Qy        783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1594 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 1653

Qy        843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1654 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 1713

Qy        903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 1773

Qy        963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1774 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1833

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1834 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1893

Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1894 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1953

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1954 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 2013

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2014 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 2073

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2074 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 2133

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2134 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 2193

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2194 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 2253

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2254 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 2313

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2314 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 2373

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2374 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2433

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2434 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 2493

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2494 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 2553

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2554 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 2613

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2614 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 2673

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 2733

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2793

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2853

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2854 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2913
```

```
Qy       2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2914 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2973

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||
Db       2974 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 3003


RESULT 13
ACN39601
ID   ACN39601 standard; cDNA; 3055 BP.
XX
AC   ACN39601;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) cDNA DNA325803, SEQ ID NO:3853.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   P-PSDB; ABM81496.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 1; SEQ ID NO 3853; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT nucleic acid of the invention
XX
SQ   Sequence 3055 BP; 736 A; 744 C; 874 G; 701 T; 0 U; 0 Other;

  Query Match             97.1%;  Score 2128.4;  DB 13;  Length 3055;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2129;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 933

Qy        123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 993

Qy        183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        994 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 1053

Qy        243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 1113

Qy        303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1114 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 1173

Qy        363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 1233

Qy        423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 1293

Qy        483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 1353

Qy        543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 1413

Qy        603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 1473

Qy        663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1533

Qy        723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1534 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 1593

Qy        783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1594 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 1653

Qy        843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1654 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 1713

Qy        903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 1773

Qy        963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1774 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1833

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1834 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1893
```

```
Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1894 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1953

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1954 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 2013

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2014 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 2073

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2074 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 2133

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2134 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 2193

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2194 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 2253

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2254 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 2313

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2314 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 2373

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2374 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2433

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2434 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 2493

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2494 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 2553

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2554 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 2613

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2614 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 2673

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 2733

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2793

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2853

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2854 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2913

Qy       2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2914 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2973

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
                 ||||||||||||||||||||||||||||||
Db       2974 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 3003
```

RESULT 14
ABL69377
ID ABL69377 standard; DNA; 2309 BP.
XX
AC ABL69377;
XX
DT 15-MAY-2002 (first entry)
XX
DE Prostate cancer related gene sequence SEQ ID NO:7714.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 7714; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2309 BP; 624 A; 571 C; 592 G; 522 T; 0 U; 0 Other;

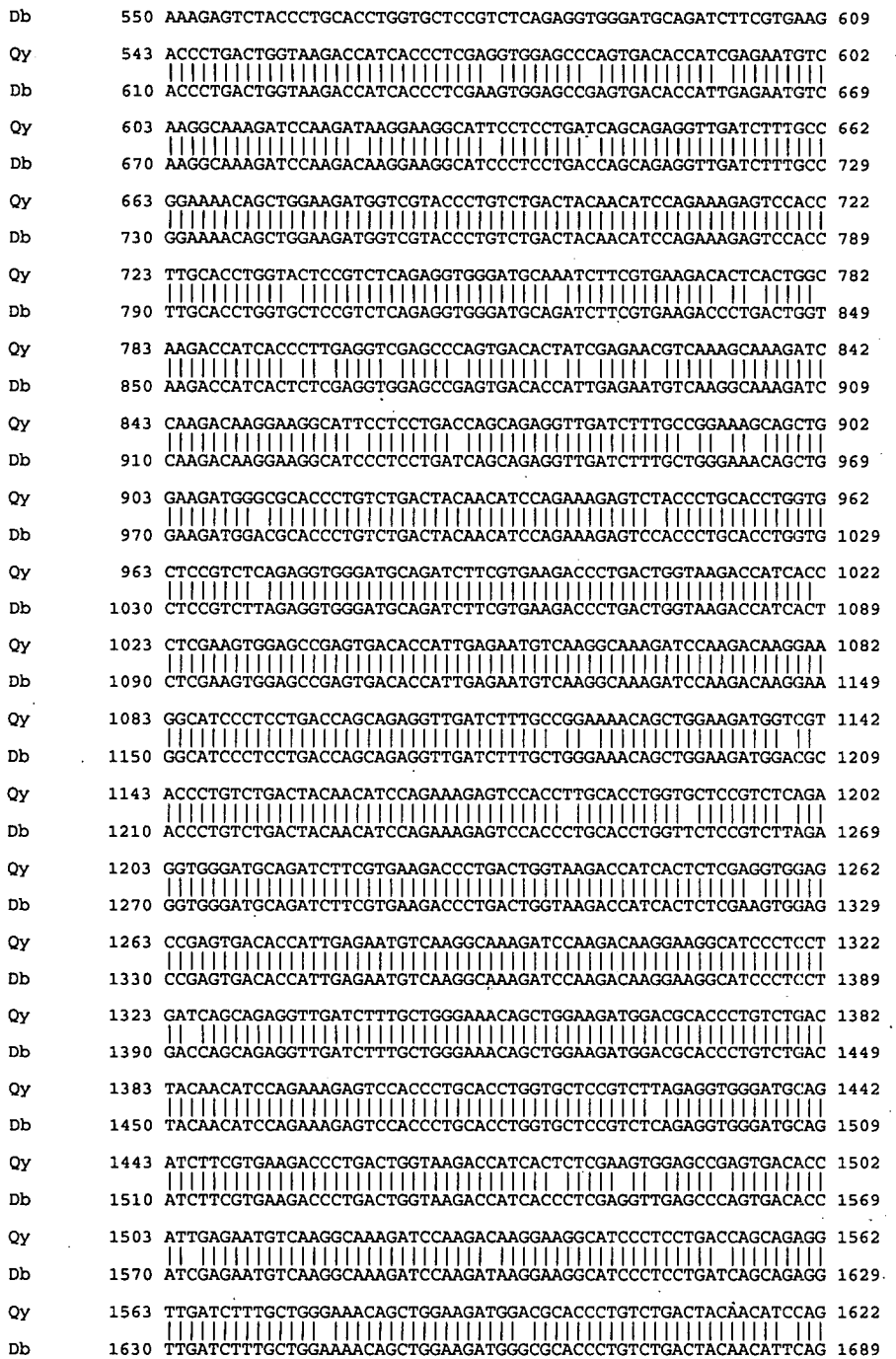
Query Match 91.2%; Score 1998.8; DB 6; Length 2309;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 2048; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy        63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
             | |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       130 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 189

Qy       123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       190 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCT 249

Qy       183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
             || |||||||| |||||||||| |||||||||||||||||||| || ||||||||||||
Db       250 GATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 309

Qy       243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
             |||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||||
Db       310 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 369

Qy       303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       370 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACA 429

Qy       363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
             |||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||
Db       430 ATCGAGAACGTCAAGGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 489

Qy       423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       490 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 549

Qy       483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       550 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 609

Qy       543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
             |||||||||||||||||||||||||||| |||||||| |||||||||| ||||||||||
Db       610 ACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 669

Qy       603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
             |||||||||||||||| |||||||||| |||||||| ||||||||||||||||||||||
Db       670 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCC 729

Qy       663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       730 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 789

Qy       723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
             |||||||||| |||||||||||||||||||||||||||| || ||||||||||||| |||||
Db       790 TTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 849

Qy       783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
             |||||||||| || ||||| ||||| |||||||| || ||||| ||||| |||||||||
Db       850 AAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 909

Qy       843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
             |||||||||||||||| |||||||| |||||||||||||||||||| || || ||||||
Db       910 CAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 969

Qy       903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
             |||||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       970 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1029

Qy       963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1030 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1089

Qy      1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1090 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1149

Qy      1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
             |||||||||||||||||||||||||||||||||||| || |||||||||||||||||| ||
Db      1150 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1209

Qy      1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
             |||||||||||||||||||||||||||||||||||| ||||||||| |||||||| |||
Db      1210 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTTCTCCGTCTTAGA 1269

Qy      1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      1270 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 1329

Qy      1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1330 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1389

Qy      1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1390 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1449

Qy      1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      1450 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1509

Qy      1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
             |||||||||||||||||||||||||||||||||||| ||||| || ||||| |||||||||
Db      1510 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 1569

Qy      1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
             || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1570 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 1629

Qy      1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||| |||
Db      1630 TTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATTCAG 1689

Qy      1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
             |||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db      1690 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 1749

Qy      1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
             || || ||||| |||||||||||| || || |||||||| |||||||||| ||||| |||
Db      1750 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 1809
```

```
Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
             ||||||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db       1810 AAGGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 1869

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
             || || |||||||||||||| || |||||||||||||||||||||||||||||||||| |||
Db       1870 GGAAAGCAGCTGGAAGATGGGCGGACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 1929

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1930 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1989

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1990 AAGACCATTACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2049

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2109

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2110 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2169

Qy       2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2170 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2229

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||||
Db       2230 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2259
```

RESULT 15
ABK83973
ID ABK83973 standard; cDNA; 2309 BP.
XX
AC ABK83973;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #544.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 544; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2309 BP; 624 A; 571 C; 592 G; 522 T; 0 U; 0 Other;

```
  Query Match            91.2%;  Score 1998.8;  DB 6;  Length 2309;
  Best Local Similarity  96.2%;  Pred. No. 0;
  Matches 2048;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy         63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
              | |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        130 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 189

Qy        123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        190 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCT 249

Qy        183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
              || ||||||||| |||||||||| |||||||||||||||||||| || ||||||||||||
Db        250 GATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 309

Qy        243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
              |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||||
Db        310 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 369

Qy        303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        370 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACA 429

Qy        363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
              |||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||||
Db        430 ATCGAGAACGTCAAGGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 489

Qy        423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 549

Qy        483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 550 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 609
Qy 543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
Db 610 ACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 669
Qy 603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
Db 670 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCC 729
Qy 663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
Db 730 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 789
Qy 723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
Db 790 TTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 849
Qy 783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
Db 850 AAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 909
Qy 843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
Db 910 CAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 969
Qy 903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
Db 970 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1029
Qy 963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
Db 1030 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1089
Qy 1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
Db 1090 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1149
Qy 1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
Db 1150 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1209
Qy 1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
Db 1210 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTTCTCCGTCTTAGA 1269
Qy 1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
Db 1270 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 1329
Qy 1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
Db 1330 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1389
Qy 1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
Db 1390 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1449
Qy 1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
Db 1450 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1509
Qy 1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
Db 1510 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 1569
Qy 1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
Db 1570 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 1629
Qy 1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
Db 1630 TTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATTCAG 1689
Qy 1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
Db 1690 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 1749
Qy 1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
Db 1750 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 1809
Qy 1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
Db 1810 AAGGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 1869
Qy 1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
Db 1870 GGAAAGCAGCTGGAAGATGGGCGGACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 1929
Qy 1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
Db 1930 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1989
Qy 1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
Db 1990 AAGACCATTACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2049
Qy 1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
Db 2050 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2109
Qy 2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
Db 2110 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2169
Qy 2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
Db 2170 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2229
Qy 2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db 2230 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2259

Search completed: March 6, 2006, 21:26:57
Job time : 1228 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 19:46:28 ; Search time 1591 Seconds
(without alignments)
11393.134 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2192	100.0	2192	6	US-10-170-385-378	Sequence 378, App
	2	2192	100.0	2192	8	US-10-684-422-1	Sequence 1, Appli
	3	2178.4	99.4	2210	8	US-10-769-476-2	Sequence 2, Appli
	4	2177	99.3	2201	8	US-10-769-476-3	Sequence 3, Appli
	5	2128.4	97.1	3055	9	US-10-956-157-4989	Sequence 4989, Ap
	6	1998.8	91.2	2309	3	US-09-969-708-243	Sequence 243, App
	7	1998.8	91.2	2309	9	US-10-843-641A-7714	Sequence 7714, Ap
	8	1691.2	77.2	2545	7	US-10-152-319A-1699	Sequence 1699, Ap
c	9	1466.8	66.9	5309	10	US-11-097-143-28699	Sequence 28699, A
c	10	1466.8	66.9	5774	10	US-11-097-143-28702	Sequence 28702, A
c	11	1466.8	66.9	5795	10	US-11-097-143-24319	Sequence 24319, A
	12	1348	61.5	1400	9	US-10-956-157-10224	Sequence 10224, A
	13	1264.8	57.7	2230	8	US-10-425-115-61463	Sequence 61463, A
	14	1250.2	57.0	2121	7	US-10-425-114-14968	Sequence 14968, A
	15	1230.4	56.1	2419	10	US-11-097-143-28700	Sequence 28700, A
	16	1230.4	56.1	2428	10	US-11-097-143-24320	Sequence 24320, A
c	17	1150.2	52.5	2378	7	US-10-424-599-2726	Sequence 2726, Ap
	18	1147.2	52.3	2067	7	US-10-437-963-48998	Sequence 48998, A
	19	1146.2	52.3	1793	7	US-10-425-114-31801	Sequence 31801, A
	20	1144.6	52.2	1893	7	US-10-425-114-33329	Sequence 33329, A
	21	1143	52.1	1790	7	US-10-425-114-30706	Sequence 30706, A
	22	1143	52.1	1885	7	US-10-425-114-31251	Sequence 31251, A
	23	1143	52.1	1896	7	US-10-425-114-32535	Sequence 32535, A
	24	1143	52.1	1902	7	US-10-425-114-33285	Sequence 33285, A
	25	1143	52.1	2110	7	US-10-425-114-5707	Sequence 5707, Ap
	26	1142.6	52.1	1956	8	US-10-425-115-177911	Sequence 177911,
	27	1140.4	52.0	1852	7	US-10-425-114-20297	Sequence 20297, A
	28	1140.4	52.0	1909	8	US-10-425-115-177909	Sequence 177909,
	29	1138.8	52.0	1813	7	US-10-425-114-32343	Sequence 32343, A
	30	1138.8	52.0	1846	7	US-10-425-114-30872	Sequence 30872, A
	31	1138.8	52.0	1846	7	US-10-425-114-30897	Sequence 30897, A
	32	1138.8	52.0	1849	7	US-10-425-114-21413	Sequence 21413, A
	33	1138.8	52.0	1850	7	US-10-425-114-30387	Sequence 30387, A
	34	1138.8	52.0	1854	7	US-10-425-114-32738	Sequence 32738, A
	35	1138.8	52.0	1854	7	US-10-425-114-33260	Sequence 33260, A
	36	1138.8	52.0	1857	7	US-10-425-114-34473	Sequence 34473, A
	37	1138.8	52.0	1858	7	US-10-425-114-5661	Sequence 5661, Ap
	38	1138.8	52.0	1858	7	US-10-425-114-18394	Sequence 18394, A
	39	1138.8	52.0	1876	7	US-10-425-114-30304	Sequence 30304, A
	40	1138.8	52.0	1879	7	US-10-425-114-5778	Sequence 5778, Ap
	41	1138	51.9	1799	7	US-10-425-114-33449	Sequence 33449, A
	42	1137.2	51.9	1836	7	US-10-425-114-4265	Sequence 4265, Ap
	43	1137.2	51.9	1838	7	US-10-425-114-4105	Sequence 4105, Ap
	44	1137.2	51.9	1868	7	US-10-425-114-20800	Sequence 20800, A
	45	1137.2	51.9	1870	7	US-10-425-114-5614	Sequence 5614, Ap

ALIGNMENTS

RESULT 1
US-10-170-385-378
; Sequence 378, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 378
; LENGTH: 2192
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-378

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
```

```
Qy       181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy       241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy       301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
            ||||||||||||||||||||||||||||||||
Db      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 2
US-10-684-422-1
; Sequence 1, Application US/10684422
; Publication No. US20040229233A1
; GENERAL INFORMATION:
; APPLICANT: ABURATANI, Hiroyuki
; APPLICANT: YAMAMOTO, Shogo
; TITLE OF INVENTION: Human housekeeping genes and human tissue-specific genes

```
;  FILE REFERENCE: 113991
;  CURRENT APPLICATION NUMBER: US/10/684,422
;  CURRENT FILING DATE:  2002-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,614
;  PRIOR FILING DATE: 2002-10-16
;  NUMBER OF SEQ ID NOS: 332
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 2192
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (69)..(2126)
US-10-684-422-1

  Query Match             100.0%;  Score 2192;  DB 8;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 3
US-10-769-476-2
; Sequence 2, Application US/10769476
; Publication No. US20050014168A1
; GENERAL INFORMATION:
;  APPLICANT: Arcturus Bioscience, Inc.
;  APPLICANT:  Erlander, Mark G.
;  APPLICANT:  Ma, Xiao-Jun
;  APPLICANT:  Baer, Thomas M.
;  TITLE OF INVENTION: 3' Biased Microarrays
;  FILE REFERENCE: 022041-001210US
;  CURRENT APPLICATION NUMBER: US/10/769,476
;  CURRENT FILING DATE:  2004-01-30
;  PRIOR APPLICATION NUMBER: US 60/475,812
;  PRIOR FILING DATE: 2003-06-03
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 2210
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-769-476-2

  Query Match             99.4%;  Score 2178.4;  DB 8;  Length 2210;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2179;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         13 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 72
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 AGGCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 66

Qy         73 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 126

Qy        133 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 186

Qy        193 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 246

Qy        253 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 306

Qy        313 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 366

Qy        373 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 426

Qy        433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 486

Qy        493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 546

Qy        553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 606

Qy        613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 666

Qy        673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 726

Qy        733 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 786

Qy        793 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 846

Qy        853 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 906

Qy        913 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCA 966

Qy        973 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGG 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGG 1026

Qy       1033 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1086

Qy       1093 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTG 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTG 1146

Qy       1153 ACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 ACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1206

Qy       1213 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACA 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACA 1266

Qy       1273 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGA 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGA 1326

Qy       1333 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1386

Qy       1393 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1446
```

```
Qy       1453 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1506

Qy       1513 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1566

Qy       1573 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1632
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1626

Qy       1633 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1686

Qy       1693 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1752
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1746

Qy       1753 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1812
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1806

Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1866

Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1926

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1986

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2052
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2046

Qy       2053 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2106

Qy       2113 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2172
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2107 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2166

Qy       2173 AATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||
Db       2167 AATAAAGTTGTTGCATTCCC 2186
```

RESULT 4
US-10-769-476-3
; Sequence 3, Application US/10769476
; Publication No. US20050014168A1
; GENERAL INFORMATION:
; APPLICANT: Arcturus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; APPLICANT: Ma, Xiao-Jun
; APPLICANT: Baer, Thomas M.
; TITLE OF INVENTION: 3' Biased Microarrays
; FILE REFERENCE: 022041-001210US
; CURRENT APPLICATION NUMBER: US/10/769,476
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/475,812
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-769-476-3

Query Match 99.3%; Score 2177; DB 8; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
              |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 5
US-10-956-157-4989
; Sequence 4989, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4989
; LENGTH: 3055
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-4989

Query Match 97.1%; Score 2128.4; DB 9; Length 3055;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2129; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

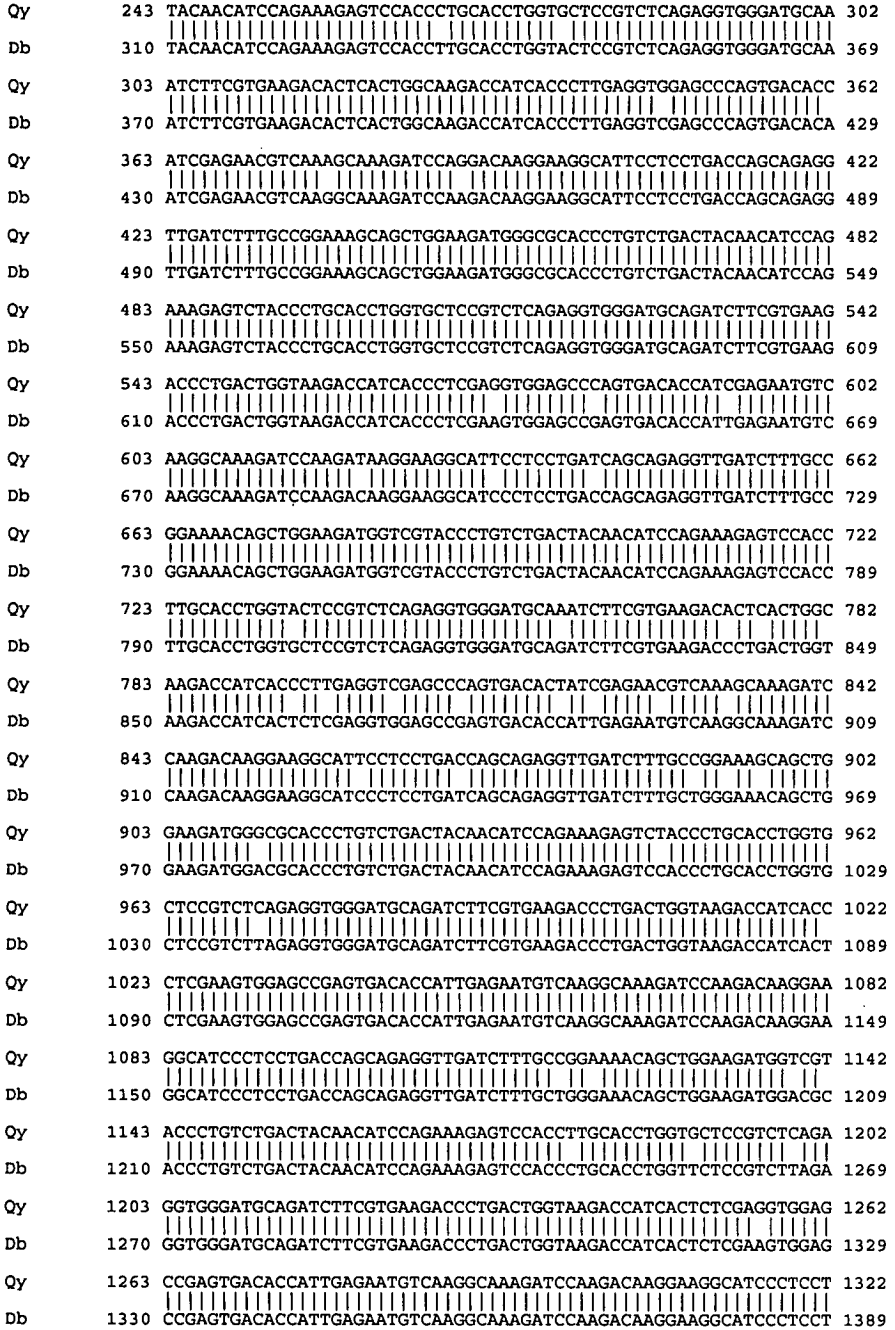
```
Qy         63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 933

Qy        123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 993

Qy        183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        994 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 1053

Qy        243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 1113

Qy        303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1114 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 1173

Qy        363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 1233

Qy        423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 1293

Qy        483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 1353

Qy        543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 1413

Qy        603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 1473
```

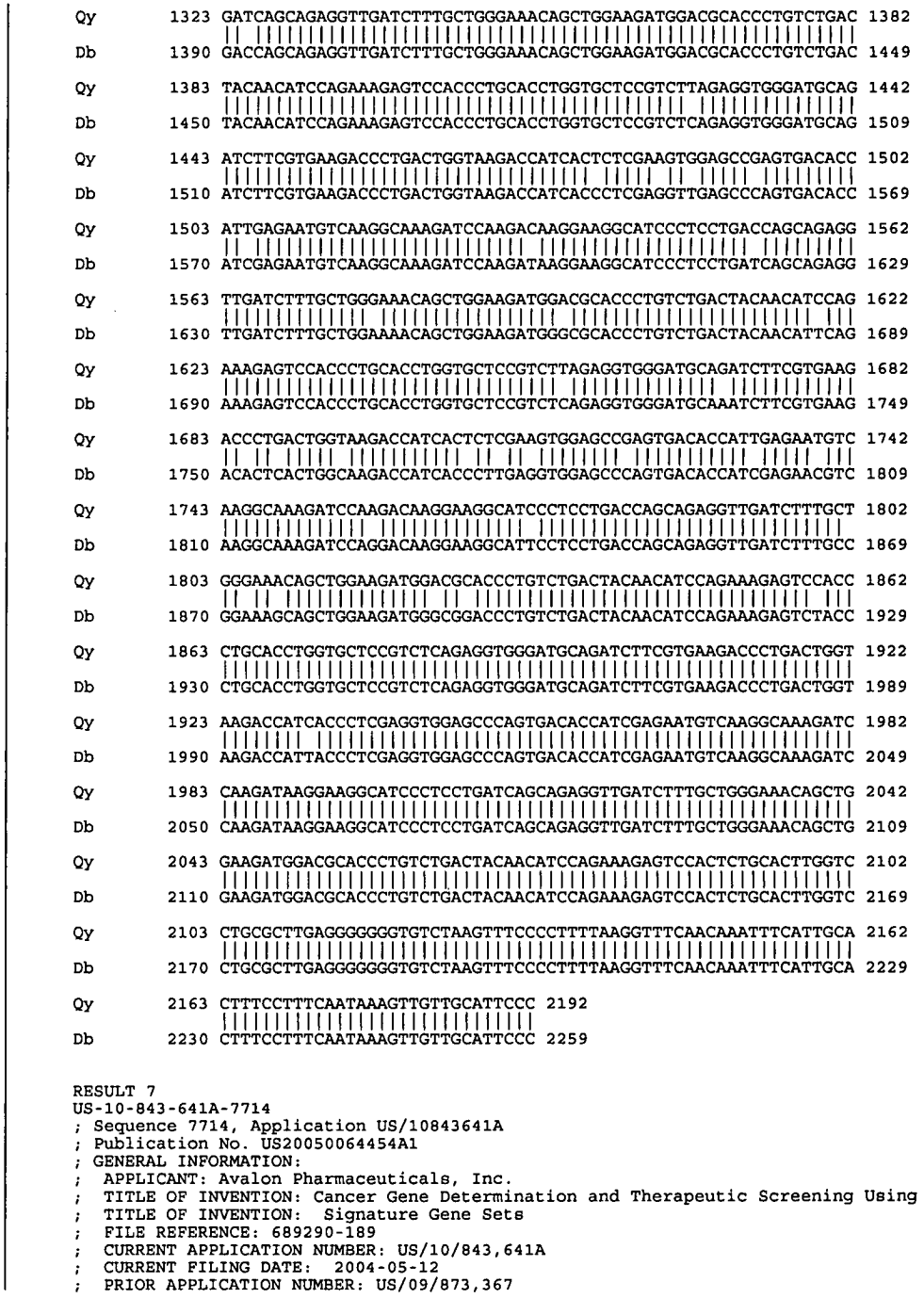
```
Qy       663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
Db      1474 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1533

Qy       723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
Db      1534 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 1593

Qy       783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
Db      1594 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 1653

Qy       843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
Db      1654 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 1713

Qy       903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
Db      1714 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 1773

Qy       963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
Db      1774 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1833

Qy      1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
Db      1834 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1893

Qy      1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
Db      1894 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1953

Qy      1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
Db      1954 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 2013

Qy      1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
Db      2014 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 2073

Qy      1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
Db      2074 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 2133

Qy      1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
Db      2134 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 2193

Qy      1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
Db      2194 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 2253

Qy      1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
Db      2254 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 2313

Qy      1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
Db      2314 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 2373

Qy      1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
Db      2374 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2433

Qy      1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
Db      2434 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 2493

Qy      1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
Db      2494 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 2553

Qy      1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
Db      2554 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 2613

Qy      1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
Db      2614 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 2673

Qy      1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
Db      2674 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 2733

Qy      1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
Db      2734 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2793

Qy      1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
Db      2794 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2853

Qy      2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
Db      2854 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2913

Qy      2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
Db      2914 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2973

Qy      2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db      2974 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 3003
```

RESULT 6
US-09-969-708-243
; Sequence 243, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 243
; LENGTH: 2309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-243

Query Match 91.2%; Score 1998.8; DB 3; Length 2309;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 2048; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy        63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
Db       130 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 189

Qy       123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
Db       190 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCT 249

Qy       183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
Db       250 GATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 309
```





RESULT 7
US-10-843-641A-7714
; Sequence 7714, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367

```
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7714
;   LENGTH: 2309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-7714

  Query Match             91.2%;  Score 1998.8;  DB 9;  Length 2309;
  Best Local Similarity   96.2%;  Pred. No. 0;
  Matches 2048;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy         63 TTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAG 122
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        130 TAGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 189

Qy        123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        190 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCT 249

Qy        183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
              || ||||||||| |||||||||| |||||||||||||||||||| || ||||||||||||
Db        250 GATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 309

Qy        243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
              |||||||||||||||||||| |||||||||| |||||||||||||||||||||||||||
Db        310 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 369

Qy        303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        370 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACA 429

Qy        363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
              |||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||
Db        430 ATCGAGAACGTCAAGGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 489

Qy        423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 549

Qy        483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 609

Qy        543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
              |||||||||||||||||||||||||||||| |||||||| |||||||||| |||||||||
Db        610 ACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 669

Qy        603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
              |||||||||||||||| |||||||||| |||||||| ||||||||||||||||||||||
Db        670 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCC 729

Qy        663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 789

Qy        723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
              |||||||||||| |||||||||||||||||||||||| |||||||||||||| || |||||
Db        790 TTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 849

Qy        783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
              |||||||||||| || ||||| ||||| |||||||| || ||||| ||||| |||||||||
Db        850 AAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 909

Qy        843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
              |||||||||||||||| |||||||| |||||||||||||||||||| || || ||||||
Db        910 CAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 969

Qy        903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 962
              |||||||| |||||||||||||||||||||||||||||||| ||||||||||||||||
Db        970 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1029

Qy        963 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1022
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1089

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1149

Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
              |||||||||||||||||||||||||||||||| || |||||||||||||||| ||
Db       1150 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1209

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
              |||||||||||||||||||||||||||||||| ||||||||| |||||||| |||
Db       1210 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTTCTCCGTCTTAGA 1269

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1270 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 1329

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1389

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1449

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1450 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1509

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
              |||||||||||||||||||||||||||||||| ||||| || ||||| |||||||||
Db       1510 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 1569

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
              || ||||||||||||||||||||||||| |||||||||||||||||||| |||||||||
Db       1570 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 1629

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
              |||||||||||||| |||||||||||||||| |||||||||||||||||||||||| |||
Db       1630 TTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATTCAG 1689

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| |||||||||
Db       1690 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 1749

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
              || || ||||| |||||||||| || || |||||||| |||||||||| ||||| |||
Db       1750 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 1809

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
              |||||||||||||| |||||||||||||| |||||||||||||||||||||||||||||
```

```
Db       1810 AAGGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 1869

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
              || || |||||||||||||| || ||||||||||||||||||||||||||||||| |||
Db       1870 GGAAAGCAGCTGGAAGATGGGCGGACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 1929

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1930 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1989

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1990 AAGACCATTACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2049

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2109

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2110 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2169

Qy       2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2170 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2229

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||
Db       2230 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2259


RESULT 8
US-10-152-319A-1699
; Sequence 1699, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Higgs, Brandon
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5089-US
;  CURRENT APPLICATION NUMBER: US/10/152,319A
;  CURRENT FILING DATE:  2002-05-22
;  PRIOR APPLICATION NUMBER: US 60/292,335
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/297,523
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,925
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,810
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/303,807
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/303,808
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: US 60/315,047
;  PRIOR FILING DATE: 2001-08-28
;  PRIOR APPLICATION NUMBER: US 60/324,928
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US 60/330,867
;  PRIOR FILING DATE: 2001-11-01
;  PRIOR APPLICATION NUMBER: US 60/330,462
;  PRIOR FILING DATE: 2001-10-22
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2221
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1699
;   LENGTH: 2545
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. NM_017314
US-10-152-319A-1699

  Query Match             77.2%;  Score 1691.2;  DB 7;  Length 2545;
  Best Local Similarity   88.9%;  Pred. No. 0;
  Matches 1828;  Conservative    0;  Mismatches  228;  Indels    0;  Gaps    0;

Qy         64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
              |||||||||||||||| ||||| ||  | |||||||||||||||||||| ||||| ||||
Db         29 TGACAATGCAGATCTTTGTGAAAACCTTAACTGGTAAGACCATCACCCTGGAGGTCGAGC 88

Qy        124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
              |||||||||||||| || |||||||||||||||||||| || ||||| |||||||| ||||
Db         89 CCAGTGACACCATTGAAAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTG 148

Qy        184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
              |||||||||||||||||||| || || |||||||||||||| |||||||||| ||||
Db        149 ACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACT 208

Qy        244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
              |||||||||||||| |||||||||||||||||||| || || |||||||| || ||||| |
Db        209 ACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCCTGCGCCTCAGAGGGGGCATGCAGA 268

Qy        304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
              |||| |||||||| || || |||||||||||||||||| ||||| ||||||||||||||||
Db        269 TCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCA 328

Qy        364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
              | || || ||||| |||||||||||||||||||| ||||| || ||||||||||||||||
Db        329 TTGAAAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGC 388

Qy        424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
              |||||||||| || |||||||||||||||| |||||||||| ||||||||||||||||||
Db        389 TGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGA 448

Qy        484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
              | ||||| ||| |||||||||| || || |||||||| || |||||||||| |||||||
Db        449 AGGAGTCCACCTTGCACCTGGTCCTTCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGA 508

Qy        544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
              ||||||| || |||||||||||||| ||||| |||||||||||||||| || |||||||
Db        509 CCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATTGAAAATGTCA 568

Qy        604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
              |||||||||||| || ||||| ||||| || ||||| ||||||||| |||||||||| |
Db        569 AGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAG 628

Qy        664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
              | || |||||||||||||| || |||||||| |||||||||||||||| |||||||||
Db        629 GCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAGGAGTCCACCC 688

Qy        724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 783
              |||||||||| || || |||||||| || ||||| ||||| |||||||| || || ||||
Db        689 TGCACCTGGTCCTTCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGACCCTGACAGGCA 748

Qy        784 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
              |||||||||||| |||||||||||||||||||| || || || ||||| ||||||||||
Db        749 AGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATTGAAAATGTCAAGGCAAAGATCC 808

Qy        844 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 903
              | |||||||| ||||| || |||||||||||||||| |||||||||| || ||||||||||
Db        809 AGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGG 868

Qy        904 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 963
              ||||||| |||||||||| |||||||||||||||||| ||||| |||||||||||||| |
Db        869 AAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCC 928

Qy        964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
              | || |||||||| || |||||||||| |||||||||||||| || ||||||||||||
Db        929 TGCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCC 988

Qy       1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083
```

```
Db        989 TGGAGGTCGAGCCCAGTGACACCATTGAAAATGTCAAGGCAAAGATCCAGGACAAGGAGG 1048

Qy       1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143

Db       1049 GCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCA 1108

Qy       1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203

Db       1109 CCCTGTCAGACTACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAG 1168

Qy       1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263

Db       1169 GGGGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGC 1228

Qy       1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323

Db       1229 CCAGTGACACCATTGAAAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTG 1288

Qy       1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383

Db       1289 ACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACT 1348

Qy       1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443

Db       1349 ACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGGGGCATGCAGA 1408

Qy       1444 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 1503

Db       1409 TCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCA 1468

Qy       1504 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1563

Db       1469 TTGAAAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGC 1528

Qy       1564 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1623

Db       1529 TGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGA 1588

Qy       1624 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1683

Db       1589 AGGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGA 1648

Qy       1684 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1743

Db       1649 CCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATTGAAAATGTCA 1708

Qy       1744 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1803

Db       1709 AGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAG 1768

Qy       1804 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1863

Db       1769 GCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAGGAGTCCACCC 1828

Qy       1864 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1923

Db       1829 TGCACCTGGTCCTGCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGACCCTGACAGGCA 1888

Qy       1924 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 1983

Db       1889 AGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATTGAAAATGTCAAGGCAAAGATCC 1948

Qy       1984 AAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 2043

Db       1949 AGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGG 2008

Qy       2044 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCC 2103

Db       2009 AAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCC 2068

Qy       2104 TGCGCTTGAGGGGGGG 2119

Db       2069 TGCGCCTCAGAGGGGG 2084
```

RESULT 9
US-11-097-143-28699/c
; Sequence 28699, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28699
; LENGTH: 5309
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-28699

Query Match 66.9%; Score 1466.8; DB 10; Length 5309;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 372; Indels 2; Gaps 2;

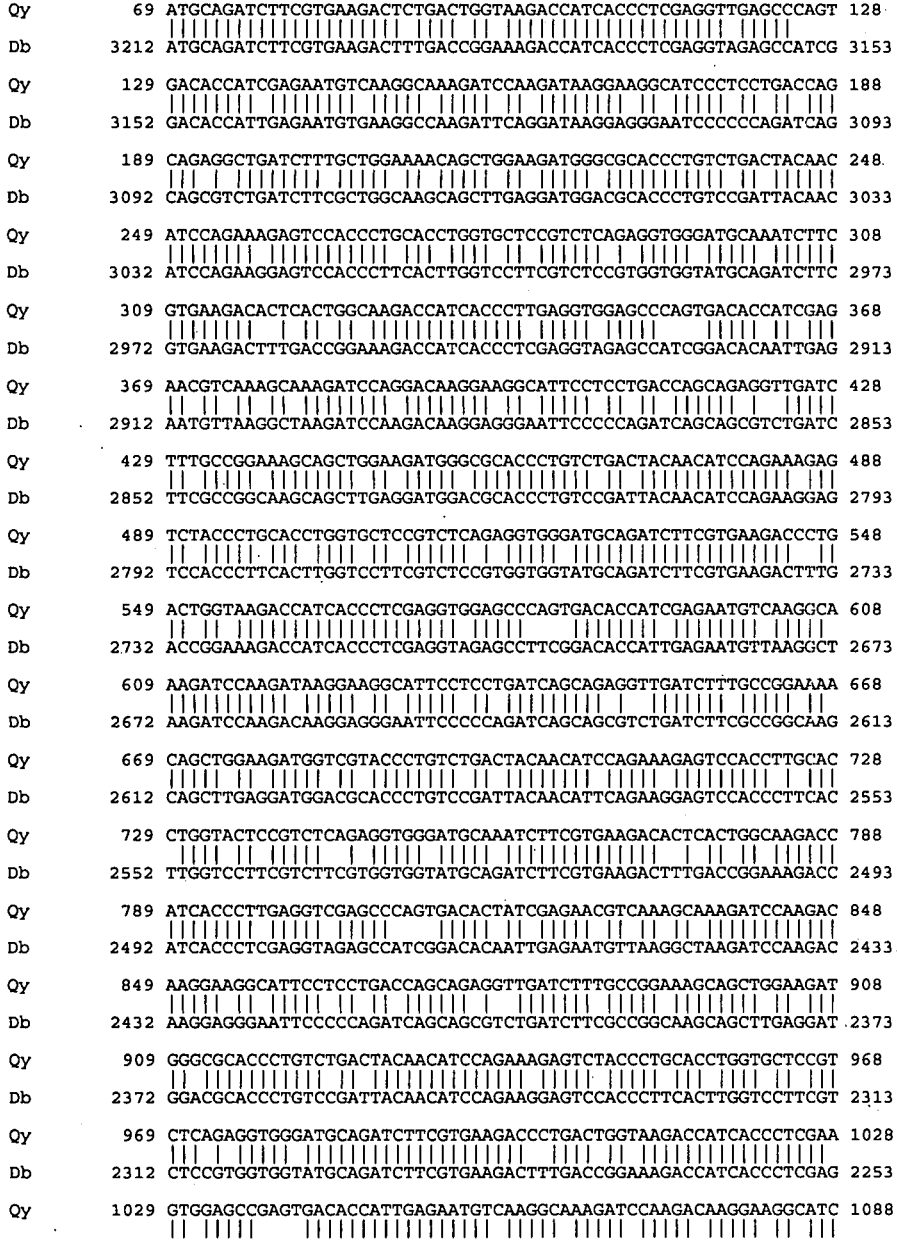
```
Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128

Db       3212 ATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCG 3153

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188

Db       3152 GACACCATTGAGAATGTGAAGGCCAAGATTCAGGATAAGGAGGGAATCCCCCCAGATCAG 3093

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248

Db       3092 CAGCGTCTGATCTTCGCTGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAAC 3033

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308

Db       3032 ATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTC 2973

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368

Db       2972 GTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCGGACACAATTGAG 2913

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428

Db       2912 AATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATC 2853

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488

Db       2852 TTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAG 2793

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
```

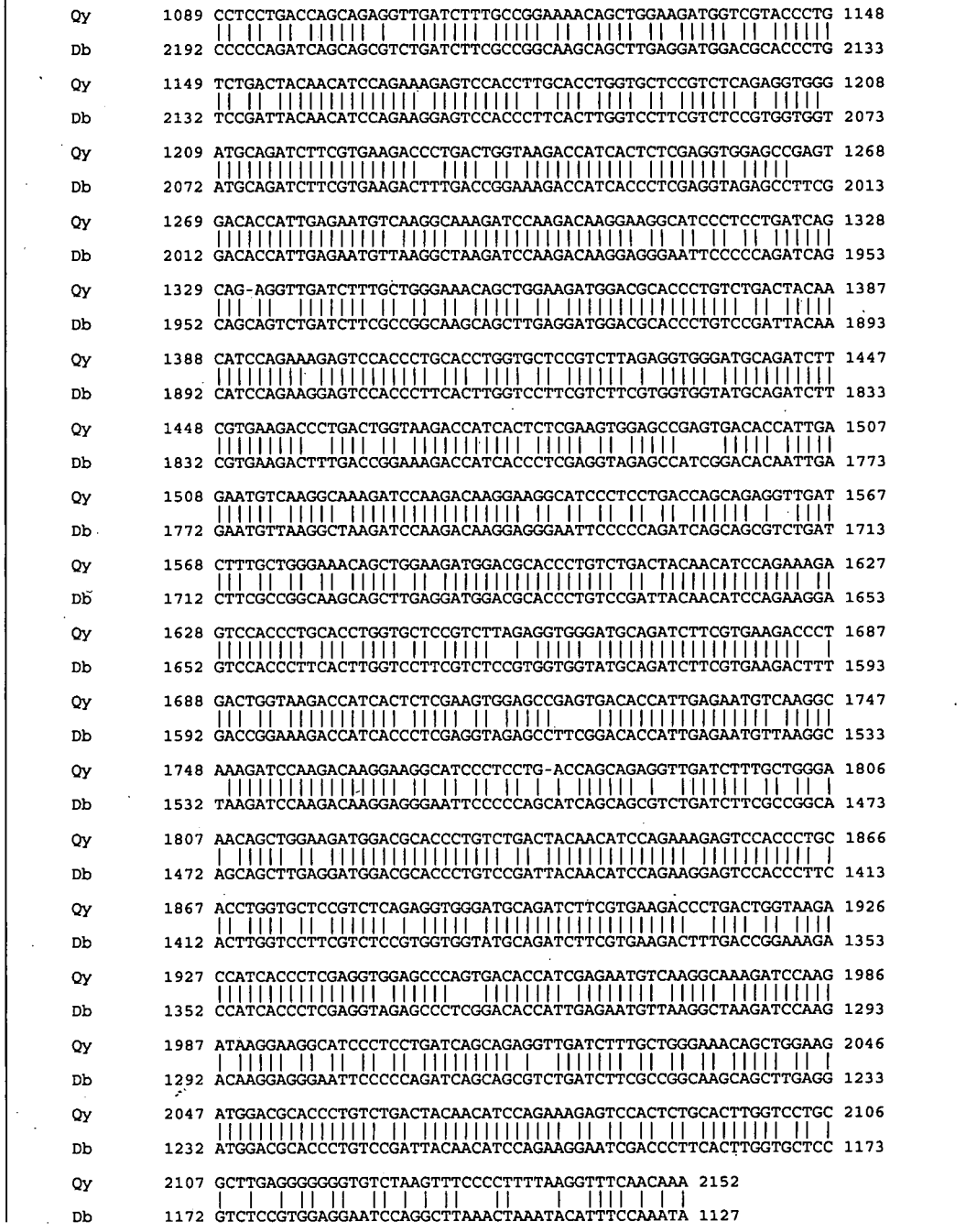
```
Db       2792 TCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTG 2733

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608

Db       2732 ACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCGGACACCATTGAGAATGTTAAGGCT 2673

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668

Db       2672 AAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAG 2613

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728

Db       2612 CAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATTCAGAAGGAGTCCACCCTTCAC 2553

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788

Db       2552 TTGGTCCTTCGTCTTCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACC 2493

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848

Db       2492 ATCACCCTCGAGGTAGAGCCATCGGACACAATTGAGAATGTTAAGGCTAAGATCCAAGAC 2433

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908

Db       2432 AAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGAT 2373

Qy        909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968

Db       2372 GGACGCACCCTGTCCGATTACAACATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGT 2313

Qy        969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028

Db       2312 CTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAG 2253

Qy       1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088

Db       2252 GTAGAGCCATCGGACACCATTGAGAATGTGAAGGCCAAGATTCAAGATAAGGAGGGAATC 2193

Qy       1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148

Db       2192 CCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTG 2133

Qy       1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208

Db       2132 TCCGATTACAACATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGT 2073

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268

Db       2072 ATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCG 2013

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328

Db       2012 GACACCATTGAGAATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAG 1953

Qy       1329 CAG-AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1387

Db       1952 CAGCAGTCTGATCTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAA 1893

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447

Db       1892 CATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTTCGTGGTGGTATGCAGATCTT 1833

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507

Db       1832 CGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCGGACACAATTGA 1773

Qy       1508 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1567

Db       1772 GAATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGAT 1713

Qy       1568 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1627

Db       1712 CTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGA 1653

Qy       1628 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1687

Db       1652 GTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTT 1593

Qy       1688 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1747

Db       1592 GACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCGGACACCATTGAGAATGTTAAGGC 1533

Qy       1748 AAAGATCCAAGACAAGGAAGGCATCCCTCCTG-ACCAGCAGAGGTTGATCTTTGCTGGGA 1806

Db       1532 TAAGATCCAAGACAAGGAGGGAATTCCCCCAGCATCAGCAGCGTCTGATCTTCGCCGGCA 1473

Qy       1807 AACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 1866

Db       1472 AGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAGTCCACCCTTC 1413

Qy       1867 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1926

Db       1412 ACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGA 1353

Qy       1927 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 1986

Db       1352 CCATCACCCTCGAGGTAGAGCCCTCGGACACCATTGAGAATGTTAAGGCTAAGATCCAAG 1293

Qy       1987 ATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 2046

Db       1292 ACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGG 1233

Qy       2047 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGC 2106

Db       1232 ATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAATCGACCCTTCACTTGGTGCTCC 1173

Qy       2107 GCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152

Db       1172 GTCTCCGTGGAGGAATCCAGGCTTAAACTAAATACATTTCCAAATA 1127
```

```
RESULT 10
US-11-097-143-28702/c
; Sequence 28702, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28702
;   LENGTH: 5774
;   TYPE: DNA
```

; ORGANISM: DROSOPHILA
US-11-097-143-28702

Query Match 66.9%; Score 1466.8; DB 10; Length 5774;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 372; Indels 2; Gaps 2;





```
RESULT 11
US-11-097-143-24319/c
; Sequence 24319, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24319
;   LENGTH: 5795
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-24319

  Query Match             66.9%;  Score 1466.8;  DB 10;  Length 5795;
  Best Local Similarity   82.1%;  Pred. No. 0;
  Matches 1712;  Conservative    0;  Mismatches  372;  Indels    2;  Gaps    2;

Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
              ||||||||||||||||||||| |||| || ||||||||||||||||||||| |||||
Db       3212 ATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCG 3153

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
              |||||||| |||||||| ||||| ||||| || |||||||| || ||||| || || |||
Db       3152 GACACCATTGAGAATGTGAAGGCCAAGATTCAGGATAAGGAGGGAATCCCCCCAGATCAG 3093

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
              ||| | |||||||| ||||| || ||||| || ||||| |||||||||||| || ||||||
Db       3092 CAGCGTCTGATCTTCGCTGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAAC 3033

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
              |||||||| |||||||||| ||| |||| || |||||| | ||||| ||||| ||||||
Db       3032 ATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTC 2973

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
              ||||||||  | || || |||||||||||||| ||||| |||||    ||||| || |||
Db       2972 GTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCGGACACAATTGAG 2913

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
              || || || || |||||||| |||||||| || ||||| || || |||||| |  |||||
Db       2912 AATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATC 2853

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
              || ||||| |||||||| || ||||| |||||||||| || |||||||||||||| |||
Db       2852 TTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAG 2793

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
              || ||||| ||| |||| || |||||| | ||||| ||||||||||||||||||||  ||

Db       2792 TCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTG 2733

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
              || || |||||||||||||||||||| |||||    ||||||| |||||||| |||||
Db       2732 ACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCGGACACCATTGAGAATGTTAAGGCT 2673

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
              |||||||||| ||||| || ||||| || |||||||| |  ||||||| ||||| ||
Db       2672 AAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAG 2613

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
              ||||| || ||||| || |||||||| || |||||||| ||||| |||||||| | |||
Db       2612 CAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATTCAGAAGGAGTCCACCCTTCAC 2553

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
              |||| || |||||  | ||||| ||||| ||||||||||||||  | || || ||||||
Db       2552 TTGGTCCTTCGTCTTCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACC 2493

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              |||||||| ||||| |||||    ||||| || ||||| || || || ||||||||||||
Db       2492 ATCACCCTCGAGGTAGAGCCATCGGACACAATTGAGAATGTTAAGGCTAAGATCCAAGAC 2433

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
              ||||| || ||||| || || |||||| |  ||||||| ||||| |||||||| || |||
Db       2432 AAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGAT 2373

Qy        909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
              || |||||||||| || |||||||||||||| ||||| ||||| ||| |||| || |||
Db       2372 GGACGCACCCTGTCCGATTACAACATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGT 2313

Qy        969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028
              ||| | ||||| |||||||||||||||||||| |||| || ||||||||||||||||||
Db       2312 CTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAG 2253

Qy       1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
              || |||||    |||||||||||||||||| ||||| ||||| ||||| |||| || |||
Db       2252 GTAGAGCCATCGGACACCATTGAGAATGTGAAGGCCAAGATTCAAGATAAGGAGGGAATC 2193

Qy       1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
              || || || |||||| |  ||||||| ||||| || ||||| || ||||| || |||||||
Db       2192 CCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTG 2133

Qy       1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
              || || |||||||||||||||||| |||||||| | ||| |||| || |||||| | |||||
Db       2132 TCCGATTACAACATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGT 2073

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
              ||||||||||||||||||||  |||| || |||||||||| |||||||| |||||
Db       2072 ATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCG 2013

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
              |||||||||||||||| ||||| |||||||||||||||| || || || || ||||||
Db       2012 GACACCATTGAGAATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAG 1953

Qy       1329 CAG-AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1387
              ||| ||  ||||||| || || || ||||| || |||||||||||||||| || |||||
Db       1952 CAGCAGTCTGATCTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAA 1893

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
              ||||||||| |||||||||| ||| |||| || |||||| | ||||| ||||||||||
Db       1892 CATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTTCGTGGTGGTATGCAGATCTT 1833

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
              |||||||||  |||| || |||||||||| ||||| || |||||    ||||| |||||
Db       1832 CGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCGGACACAATTGA 1773

Qy       1508 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1567
              |||||| ||||| |||||||||||||||||||| || || || || |||||| |  ||||
Db       1772 GAATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGAT 1713

Qy       1568 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1627
              ||| || || || ||||| || |||||||||||||||| || |||||||||||||| ||
Db       1712 CTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGA 1653
```

```
Qy      1628 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1687
             ||||||||| ||| |||| || |||||  | |||||| |||||||||||||||||||||  |
Db      1652 GTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTT 1593

Qy      1688 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1747
             ||| || |||||||||||| ||||| || |||||     ||||||||||||||||||| |||||
Db      1592 GACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCGGACACCATTGAGAATGTTAAGGC 1533

Qy      1748 AAAGATCCAAGACAAGGAAGGCATCCCTCCTG-ACCAGCAGAGGTTGATCTTTGCTGGGA 1806
             |||||||||||||||| || || || || | | |||||| |  ||||||| || || |
Db      1532 TAAGATCCAAGACAAGGAGGGAATTCCCCCAGCATCAGCAGCGTCTGATCTTCGCCGGCA 1473

Qy      1807 AACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 1866
             | |||||| || |||||||||||||||||| || |||||||||||||| |||||||||| |
Db      1472 AGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAGTCCACCCTTC 1413

Qy      1867 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1926
             || |||| || |||||| | ||||| |||||||||||||||||||  |||| || ||||
Db      1412 ACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGA 1353

Qy      1927 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 1986
             ||||||||||||||| ||||||   |||||||| |||||||| ||||| ||||||||||
Db      1352 CCATCACCCTCGAGGTAGAGCCCTCGGACACCATTGAGAATGTTAAGGCTAAGATCCAAG 1293

Qy      1987 ATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 2046
             | ||||| || || || || ||||||||| |  ||||||| || || || ||||| || |
Db      1292 ACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGG 1233

Qy      2047 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGC 2106
             |||||||||||||||| || |||||||||||||| || || || || |||||||| || |
Db      1232 ATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAATCGACCCTTCACTTGGTGCTCC 1173

Qy      2107 GCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
             |  |  | || ||  || | | ||   ||     |  |||| | | |
Db      1172 GTCTCCGTGGAGGAATCCAGGCTTAAACTAAATACATTTCCAAATA 1127


RESULT 12
US-10-956-157-10224
; Sequence 10224, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10224
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-10224

  Query Match           61.5%;  Score 1348;  DB 9;  Length 1400;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1348;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 60

Qy       905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 120

Qy       965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 180

Qy      1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 240

Qy      1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 300

Qy      1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 360

Qy      1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 420

Qy      1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 480

Qy      1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 540

Qy      1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 600

Qy      1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 660

Qy      1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 720

Qy      1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 780

Qy      1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 840

Qy      1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 900

Qy      1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 960

Qy      1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1020

Qy      1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1924
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1080

Qy      1925 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1984
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1140

Qy      1985 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 1200

Qy      2045 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 1260
```

```
Qy       2105 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 1320

Qy       2165 TTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||
Db       1321 TTCCTTTCAATAAAGTTGTTGCATTCCC 1348


RESULT 13
US-10-425-115-61463
; Sequence 61463, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 61463
;   LENGTH: 2230
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_15604C.1
US-10-425-115-61463

  Query Match             57.7%;  Score 1264.8;  DB 8;  Length 2230;
  Best Local Similarity   81.3%;  Pred. No. 0;
  Matches 1505;  Conservative    0;  Mismatches  342;  Indels    5;  Gaps    3;

Qy         59 TCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGT 118
              ||   ||  ||||||||||||||| ||||| || ||||| ||||| |||||||||||||||
Db        192 TCTTCTGCAGATGCAGATCTTCGTCAAGACCCTCACTGGAAAGACTATCACCCTCGAGGT 251

Qy        119 TGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCC 178
              ||| |   ||||||||| || || |||||||| |||||||| || ||||| ||||||||
Db        252 GGAGTCTTCTGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCC 311

Qy        179 TCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTC 238
              || ||||||||| |||| |||||||||| || ||||| || || |||||||| ||   |
Db        312 CCCGGACCAGCAGCGGCTCATCTTTGCTGGCAAGCAGCTCGAGGACGGGCGCACACTTGC 371

Qy        239 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 298
              |||||||||||||||| |||  |||||| ||||| || ||||| ||||| || || ||
Db        372 CGACTACAACATCCAGAAGGAGAGCACCCTCCACCTTGTTCTCCGCCTCAGGGGAGGCAT 431

Qy        299 GCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGA 358
              ||| |||||||| ||||| |||||||| ||||| |||||||| ||||||||  |   |||
Db        432 GCAGATCTTCGTCAAGACCCTCACTGGAAAGACTATCACCCTCGAGGTGGAGTCTTCTGA 491

Qy        359 CACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCA 418
              ||||||||| |||||||| || |||||||||||||||| ||||| ||||| ||||||||
Db        492 CACCATCGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCA 551

Qy        419 GAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 478
              | || | || |||||||| |||||||| || |||||||| ||| | ||||||||||||||
Db        552 GCGGCTCATTTTTGCCGGCAAGCAGCTCGAGGACGGGCGCACGCTGGCCGACTACAACAT 611

Qy        479 CCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGT 538
              |||||  |||   ||||| ||| ||||||| |||||||| || || ||||||||||||||
Db        612 TCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGTCTCAGGGGCGGCATGCAGATCTTCGT 671

Qy        539 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 598
              ||||||  ||||| || ||||| |||||||| ||||| |||  |   ||| ||||| || ||
Db        672 GAAGACTCTGACCGGCAAGACTATCACCCTTGAGGTCGAGTCTTCTGATACCATTGATAA 731

Qy        599 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTT 658
```

```
Db        732 TGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCTCCGGACCAGCAGCGGCTCAT-TT 790

Qy        659 TGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTC 718
              |||||| || |||||||| |||||||| ||| || |||||||||||||| ||||| |||
Db        791 TGCCGGCAAGCAGCTGGAGGATGGTCGCACCTTGGCTGACTACAACATTCAGAAGGAGAG 850

Qy        719 CACCTTGCACCTGGTACTCCGTCTCAGAGGTGGG--ATGCAAATCTTCGTGAAGAC--AC 774
              |||| | ||| |||| || || ||||| ||  ||  ||||| || || ||||||||
Db        851 CACCCTCCACTTGGTGCTGCGCCTCAGGGGGTGGCTATGCAGATTTTTGTGAAGACCACT 910

Qy        775 TCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAG 834
              | |||||||||||||||||| ||||||| ||| |   |||||| || || || ||||| |
Db        911 TGACTGGCAAGACCATCACCTTTGAGGTGGAGTCGTCTGACACCATTGATAATGTCAAGG 970

Qy        835 CAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 894
              | |||||||| |||||||| ||||| ||||| |||||||||| || | || |||||||| |
Db        971 CCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCAGCGGCTCATTTTTGCCGGCA 1030

Qy        895 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 954
              ||||||| || || |||||||| ||| | |||||||||| ||||| |||   ||||| |
Db       1031 AGCAGCTCGAGGACGGGCGCACGCTGGCCGACTACAACATTCAGAAGGAGAGCACCCTCC 1090

Qy        955 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1014
              || ||||||| |||||||| || || ||||||||||||||||||| ||||| || ||||
Db       1091 ACTTGGTGCTGCGTCTCAGGGGCGGCATGCAGATCTTCGTGAAGACTCTGACCGGCAAGA 1150

Qy       1015 CCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1074
              | |||||||| || || ||| |   ||| |||||||| |||||||||| |||||||| |
Db       1151 CTATCACCCTTGAGGTCGAGTCTTCTGATACCATTGATAATGTCAAGGCCAAGATCCAGG 1210

Qy       1075 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 1134
              ||||||| |||||||||| ||||||||| || | || |||||||| || ||||| || |
Db       1211 ACAAGGAGGGCATCCCTCCGGACCAGCAGCGGCTCATTTTTGCCGGCAAGCAGCTCGAGG 1270

Qy       1135 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCC 1194
              | || || || ||| | |||||||||| ||||| |||  |||| | ||| ||||||| |
Db       1271 ACGGGCGCACGCTGGCCGACTACAACATTCAGAAGGAGAGCACCCTCCACTTGGTGCTGC 1330

Qy       1195 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 1254
              ||||||| ||||| |||||||||||||||||| || ||||| ||||| ||||| || |
Db       1331 GTCTCAGGGGTGGCATGCAGATCTTCGTGAAGACACTTACTGGCAAGACGATCACCCTTG 1390

Qy       1255 AGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 1314
              |||||||| |    |||||||||| || |||||||| |||||||| |||||||| ||||
Db       1391 AGGTGGAGTCCTCAGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCA 1450

Qy       1315 TCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCC 1374
              | || || || |||||| || ||||||| || || || |||||||| ||||| ||||||
Db       1451 TTCCACCGGACCAGCAGCGGCTGATCTTCGCCGGCAAGCAGCTGGAGGATGGTCGCACCT 1510

Qy       1375 TGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTG 1434
              || |||||||||||||||||| |||  |||||| |||||||||||||| || || ||||
Db       1511 TGGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACCTGGTGCTCCGCCTCAGGGGTG 1570

Qy       1435 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGA 1494
              | |||||||| || ||||||||  ||||||| ||||||||||  | || |||||| ||
Db       1571 GTATGCAGATTTTTGTGAAGACATTGACTGGCAAGACCATCACCTTGGAGGTGGAGTCGT 1630

Qy       1495 GTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACC 1554
              ||||||  || || |||||||| || |||||||| |||||||| ||||| || || ||||
Db       1631 CTGACACAATCGATAATGTCAAAGCCAAGATCCAGGACAAGGAGGGCATTCCACCGGACC 1690

Qy       1555 AGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACA 1614
              |||||  |||||||||||||| || |||||||| ||||| || |||||  | |||||||
Db       1691 AGCAGCGCCTGATCTTTGCTGGCAAGCAGCTGGAGGATGGCCGTACCCTCGCCGACTACA 1750

Qy       1615 ACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCT 1674
              |||||||||| |||  |||||| ||||||||||||||  |  ||||||| ||||||||||
Db       1751 ACATCCAGAAGGAGAGCACCCTCCACCTGGTGCTCCGCTTGCGAGGTGGTATGCAGATCT 1810

Qy       1675 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTG 1734
              | |||||||| |||||||  |||||||| || || || || |||      |||||||| |
```

```
Db       1811 TTGTGAAGACCTTGACTGGCAAGACCATAACCCTGGAGGTTGAGAGCTCCGACACCATCG 1870

Qy       1735 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGA 1794

Db       1871 ACAACGTGAAGGCAAAGATCCAGGACAAAGAGGGCATCCCCCCAGACCAGCAGCGTCTGA 1930

Qy       1795 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 1854

Db       1931 TCTTCGCTGGCAAGCAGCTGGAGGATGGTCGCACCCTAGCAGACTACAACATCCAGAAGG 1990

Qy       1855 AGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGT 1906

Db       1991 AGAGCACTCTCCACCTGGTGCTCCGCCTCCGTGGTGGTCAGTAAGCCATGGT 2042


RESULT 14
US-10-425-114-14968
; Sequence 14968, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14968
;   LENGTH: 2121
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3049-005-H6_FLI
US-10-425-114-14968

  Query Match             57.0%;  Score 1250.2;  DB 7;  Length 2121;
  Best Local Similarity   79.4%;  Pred. No. 0;
  Matches 1480;  Conservative    0;  Mismatches  383;  Indels    0;  Gaps    0;

Qy         54 GATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTC 113

Db         52 GCTCACCGTTTCAAGATGCAAATCTTCGTTAAGACCCTCACTGGTAAGACCATCACTCTC 111

Qy        114 GAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 173

Db        112 GAAGTCGAGAGTTCAGATACCATAGACAATGTAAAGGCCAAAATTCAAGACAAGGAAGGG 171

Qy        174 ATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACC 233

Db        172 ATCCCACCAGATCAGCAACGGCTCATTTTTGCTGGCAAGCAGCTTGAGGATGGACGTACC 231

Qy        234 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT 293

Db        232 CTTGCTGATTATAACATCCAAAAGGAGTCCACTCTCCACTTGGTTCTTCGTCTCCGTGGT 291

Qy        294 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCC 353

Db        292 GGGATGCAAATCTTTGTTAAGACCTTGACGGGAAAGACAATCACCCTTGAGGTCGAGAGC 351

Qy        354 AGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGAC 413

Db        352 TCCGACACCATTGACAACGTTAAGGCTAAGATTCAGGACAAAGAGGGTATCCCTCCAGAT 411

Qy        414 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 473

Db        412 CAACAGAGGCTGATCTTTGCTGGAAAGCAGCTTGAGGATGGCCGTACCCTTGCTGATTAC 471

Qy        474 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 533

Db        472 AACATCCAGAAGGAGTCCACTCTTCATTTGGTACTCCGTCTACGTGGTGGAATGCAAATC 531

Qy        534 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 593

Db        532 TTTGTTAAGACTCTGACCGGAAAGACAATCACCCTTGAGGTGGAGAGCTCTGACACCATT 591

Qy        594 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTG 653

Db        592 GACAATGTTAAGGCCAAGATTCAGGACAAAGAGGGTATCCCACCAGATCAACAGAGGCTC 651

Qy        654 ATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAA 713

Db        652 ATCTTTGCTGGAAAGCAGCTTGAGGATGGCCGTACCCTTGCTGATTACAACATCCAGAAG 711

Qy        714 GAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 773

Db        712 GAGTCCACTCTTCACTTGGTACTCCGTCTACGTGGTGGAATGCAAATCTTTGTTAAGACT 771

Qy        774 CTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAA 833

Db        772 CTGACCGGAAAGACAATCACCCTTGAGGTGGAGAGCTCTGACACCATTGACAATGTTAAG 831

Qy        834 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 893

Db        832 GCTAAGATTCAGGACAAAGAGGGTATCCCACCAGACCAACAGAGGCTCATCTTTGCTGGA 891

Qy        894 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 953

Db        892 AAGCAGCTAGAAGATGGCCGAACCTTGGCTGATTACAACATCCAAAAAGAATCTACCCTC 951

Qy        954 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1013

Db        952 CACTTGGTTCTCCGCCTTCGTGGTGGGATGCAAATCTTTGTCAAGACTCTCACTGGAAAG 1011

Qy       1014 ACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 1073

Db       1012 ACCATCACCCTTGAGGTGGAGAGCTCTGACACTATTGACAATGTCAAAGCCAAGATTCAG 1071

Qy       1074 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 1133

Db       1072 GACAAAGAAGGTATCCCACCAGACCAACAGAGGCTCATTTTTGCTGGCAAACAGCTTGAA 1131

Qy       1134 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTC 1193

Db       1132 GATGGCCGAACCTTGGCTGATTACAACATCCAAAAAGAATCTACCCTCCACTTGGTGCTC 1191

Qy       1194 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 1253

Db       1192 CGTCTGCGCGGTGGTATGCAAATCTTTGTGAAGACCCTTACTGGGAAGACCATTACTTTG 1251

Qy       1254 GAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1313

Db       1252 GAGGTAGAGAGCTCAGACACAATTGACAATGTGAAGGCAAAGATCCAAGACAAGGAGGGT 1311

Qy       1314 ATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 1373

Db       1312 ATCCCACCAGACCAACAGAGGCTCATTTTTGCTGGCAAACAGCTTGAAGATGGCCGAACC 1371

Qy       1374 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGT 1433

Db       1372 TTGGCTGATTACAACATCCAAAAAGAATCTACCCTCCACTTGGTGCTCCGTCTGCGCGGT 1431

Qy       1434 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 1493

Db       1432 GGTATGCAAATCTTTGTGAAGACCCTTACTGGGAAGACCATTACTTTGGAGGTAGAGAGC 1491

Qy       1494 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1553

Db       1492 TCAGACACAATTGACAATGTGAAGGCAAAGATCCAAGACAAGGAGGGTATTCCTCCAGAC 1551

Qy       1554 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1613
```

```
Db       1552 CAGCAAAGGTTGATTTTTGCTGGAAAGCAGCTTGAAGATGGACGTACACTGGCAGACTAC 1611

Qy       1614 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1673
             |||||||| || |||||||||| ||||| || || ||||| || || || |||||||||
Db       1612 AACATCCAAAAGGAGTCCACCCTTCACCTTGTCCTTCGTCTCAGGGGAGGTATGCAGATC 1671

Qy       1674 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1733
             || ||||||||  |||| || ||||| || ||  | ||||| ||        ||||| ||
Db       1672 TTTGTGAAGACTTTGACAGGCAAGACTATTACCTTGGAAGTTGAAAGCTCAGACACAATC 1731

Qy       1734 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1793
             || |||||||| ||||| ||||| |||||||| || ||||| || ||||||||||||||||
Db       1732 GATAATGTCAAAGCAAAAATCCAGGACAAGGAGGGTATCCCACCAGACCAGCAGAGGTTG 1791

Qy       1794 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1853
             || |||||||||| ||  |||| |||||| | || ||  ||||||||||||| ||||||
Db       1792 ATATTTGCTGGGAAGCAATTGGAGGATGGAAGGACTCTTGCTGACTACAACATTCAGAAA 1851

Qy       1854 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1913
             ||||| || || ||||| ||| | |||||  | ||||||  |  | |  | | |||
Db       1852 GAGTCTACTCTTCACCTTGTGTTGCGTCTTCGTGGTGGGTTTTAAAAGTCTGAGAATGTT 1911

Qy       1914 CTG 1916
              |||
Db       1912 CTG 1914


RESULT 15
US-11-097-143-28700
; Sequence 28700, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28700
;   LENGTH: 2419
;   TYPE: DNA
;   ORGANISM: DROSOPHILA
US-11-097-143-28700

  Query Match           56.1%;  Score 1230.4;  DB 10;  Length 2419;
  Best Local Similarity  76.6%;  Pred. No. 0;
  Matches 1596;  Conservative    0;  Mismatches  356;  Indels  132;  Gaps    2;

Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
              |||||||||||||||||||| |||| || ||||||||||||||||||| |||||
Db        342 ATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCG 401

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
              |||||||| |||||||| ||||| ||||| || |||||||| || ||||| || || |||
Db        402 GACACCATTGAGAATGTGAAGGCCAAGATTCAGGATAAGGAGGGAATCCCCCCAGATCAG 461

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
              ||| | |||||||| ||||| || ||||| || ||||| |||||||||| || ||||||
Db        462 CAGCGTCTGATCTTCGCTGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAAC 521

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
              |||||||| |||||||||| ||| |||| || |||||| | ||||| ||||| ||||||
Db        522 ATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTC 581

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
              ||||||||  | || || |||||||||||||| ||||| |||||    ||||| || |||
Db        582 GTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCGGACACAATTGAG 641

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
              || || || || |||||||| |||||||| || ||||| || || |||||| |  |||||
Db        642 AATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATC 701

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
              || ||||| |||||||| || ||||| |||||||||| || |||||||||||||| |||
Db        702 TTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAG 761

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
              || ||||| ||| |||| || |||||| | ||||| ||||||||||||||||||||  ||
Db        762 TCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTG 821

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
              || || |||||||||||||||||||| |||||    |||||||| |||||||| |||||
Db        822 ACCGGAAAGACCATCACCCTCGAGGTAGAGCCTTCGGACACCATTGAGAATGTTAAGGCT 881

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
              |||||||||| ||||| || ||||| || ||||||||| |  ||||||| ||||| ||
Db        882 AAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAG 941

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
              ||||| || |||||| || |||||||| || |||||||| ||||| |||||||| | |||
Db        942 CAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATTCAGAAGGAGTCCACCCTTCAC 1001

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
              |||| || |||||  | ||||| ||||| ||||||||||||||  | || || |||||||
Db       1002 TTGGTCCTTCGTCTTCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACC 1061

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              ||||||| ||||| |||||    ||||| || ||||| || || || ||||||||||||
Db       1062 ATCACCCTCGAGGTAGAGCCATCGGACACAATTGAGAATGTTAAGGCTAAGATCCAAGAC 1121

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
              ||||| || ||||| || || |||||| |  |||||| ||||| |||||||| || |||
Db.      1122 AAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGAT 1181

Qy        909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
              || |||||||||| || |||||||||||||| ||||| ||||| ||| |||| || |||
Db       1182 GGACGCACCCTGTCCGATTACAACATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGT 1241

Qy        969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028
              ||| | ||||| |||||||||||||||||||  |||| || ||||||||||||||||||
Db       1242 CTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAG 1301

Qy       1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
              || |||||    |||||||||||||||| ||||| ||||| ||||| ||||| || |||
Db       1302 GTAGAGCCATCGGACACCATTGAGAATGTGAAGGCCAAGATTCAAGATAAGGAGGGAATC 1361

Qy       1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
              || || || |||||| |  |||||| ||||| || ||||| || ||||| || ||||||
Db       1362 CCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTG 1421

Qy       1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
              || || |||||||||||||| |||||||| | ||| |||| || |||||| | |||||
Db       1422 TCCGATTACAACATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGT 1481

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
              ||||||||||||||||||||  |||| || |||||||||| |||||| |       |||
```

```
Db       1482 ATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGATCAGCAGCAGT 1541

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328

Db       1542 ------------------------------------------------------------ 1541

Qy       1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
                 ||||||| || || || ||||| || |||||||||||||||||| || ||||||
Db       1542 ------CTGATCTTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAAC 1595

Qy       1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448
              |||||||| |||||||||| ||| |||| || |||||| | ||||| ||||||||||
Db       1596 ATCCAGAAGGAGTCCACCCTTCACTTGGTCCTTCGTCTTCGTGGTGGTATGCAGATCTTC 1655

Qy       1449 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1508
              |||||||| |||| || |||||||||| ||||| || |||||     ||||| ||||||
Db       1656 GTGAAGACTTTGACCGGAAAGACCATCACCCTCGAGGTAGAGCCATCGGACACAATTGAG 1715

Qy       1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
              ||||| ||||| |||||||||||||||||| || || || || || |||||| |  |||||
Db       1716 AATGTTAAGGCTAAGATCCAAGACAAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATC 1775

Qy       1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
              || || || || ||||| || |||||||||||||||||| || |||||||||||||| |||
Db       1776 TTCGCCGGCAAGCAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAG 1835

Qy       1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
              |||||||| ||| |||| || |||||   | ||||| ||||||||||||||||||||  ||
Db       1836 TCCACCCTTCACTTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTG 1895

Qy       1689 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1748
              || || |||||||||| |||||
Db       1896 ACCGGAAAGACCATCACCCTCGAGC---------------------------------- 1920

Qy       1749 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1808
                                          | |||||| |  ||||||| || || ||
Db       1921 -------------------------------ATCAGCAGCGTCTGATCTTCGCCGGCAAG 1949

Qy       1809 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1868
              ||||| || |||||||||||||||| || |||||||||||| |||||||||| |||
Db       1950 CAGCTTGAGGATGGACGCACCCTGTCCGATTACAACATCCAGAAGGAGTCCACCCTTCAC 2009

Qy       1869 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1928
               |||| || |||||| | |||||| ||||||||||||||||||||  |||| || ||||||
Db       2010 TTGGTCCTTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACTTTGACCGGAAAGACC 2069

Qy       1929 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 1988
              |||||||||||||| ||||||   |||||||| |||||||| ||||| ||||||||||
Db       2070 ATCACCCTCGAGGTAGAGCCCTCGGACACCATTGAGAATGTTAAGGCTAAGATCCAAGAC 2129

Qy       1989 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 2048
              ||||| || || || || ||||||||| |  ||||||| || || || ||||| || |||
Db       2130 AAGGAGGGAATTCCCCCAGATCAGCAGCGTCTGATCTTCGCCGGCAAGCAGCTTGAGGAT 2189

Qy       2049 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2108
              |||||||||||||| || |||||||||||||| || || || || |||||||| || ||
Db       2190 GGACGCACCCTGTCCGATTACAACATCCAGAAGGAATCGACCCTTCACTTGGTGCTCCGT 2249

Qy       2109 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
                |  | || ||  || | | ||   ||     |  |||| | | |
Db       2250 CTCCGTGGAGGAATCCAGGCTTAAACTAAATACATTTCCAAATA 2293
```

Search completed: March 6, 2006, 20:19:23
Job time : 1598 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 19:52:44 ; Search time 657 Seconds
(without alignments)
7314.958 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2177	99.3	2201	9	US-11-050-857-1056	Sequence 1056, Ap
	2	2177	99.3	2201	12	US-11-043-590-46	Sequence 46, Appl
	3	1644.4	75.0	2764	12	US-11-136-527-2416	Sequence 2416, Ap
	4	1447.8	66.0	2346	12	US-11-102-026A-8	Sequence 8, Appli
	5	909.4	41.5	1082144	12	US-11-117-187-211	Sequence 211, App
	6	810	37.0	1400	12	US-11-136-527-6512	Sequence 6512, Ap
	7	718.4	32.8	1223	12	US-11-136-527-1813	Sequence 1813, Ap
	8	718.4	32.8	1223	12	US-11-136-527-5909	Sequence 5909, Ap
c	9	565	25.8	1499	8	US-10-750-185-45440	Sequence 45440, A
c	10	565	25.8	1499	8	US-10-750-623-45440	Sequence 45440, A
	11	552.2	25.2	3107	12	US-11-122-329-68	Sequence 68, Appl
	12	474	21.6	3083	8	US-10-927-641-1	Sequence 1, Appli
	13	474	21.6	3084	8	US-10-927-641-143	Sequence 143, App
c	14	469.4	21.4	1082144	12	US-11-117-187-211	Sequence 211, App
	15	433.4	19.8	460	12	US-11-108-172-163	Sequence 163, App
	16	398.2	18.2	412	12	US-11-108-172-660	Sequence 660, App
	17	383.6	17.5	551	12	US-11-102-026A-100	Sequence 100, App
	18	370.2	16.9	1608	12	US-11-128-061-550	Sequence 550, App
	19	370.2	16.9	1608	12	US-11-128-049-550	Sequence 550, App
	20	338.2	15.4	1274	8	US-10-927-641-34	Sequence 34, Appl
	21	262.6	12.0	407	12	US-11-102-026A-103	Sequence 103, App
	22	260.8	11.9	1920	12	US-11-120-351A-1	Sequence 1, Appli
	23	259.8	11.9	459	12	US-11-120-351A-3	Sequence 3, Appli
	24	257.2	11.7	8751	12	US-11-128-061-876	Sequence 876, App
	25	257.2	11.7	8751	12	US-11-128-049-876	Sequence 876, App
	26	224.2	10.2	535	9	US-11-116-881A-1346	Sequence 1346, Ap
	27	200.6	9.2	581	12	US-11-128-061-2181	Sequence 2181, Ap
	28	200.6	9.2	581	12	US-11-128-049-2181	Sequence 2181, Ap
	29	198.8	9.1	14002	12	US-11-119-330-6	Sequence 6, Appli
	30	196.6	9.0	2832	12	US-11-128-061-3421	Sequence 3421, Ap
	31	196.6	9.0	2832	12	US-11-128-049-3421	Sequence 3421, Ap
	32	190.2	8.7	3295	12	US-11-128-061-593	Sequence 593, App
	33	190.2	8.7	3295	12	US-11-128-049-593	Sequence 593, App
	34	185	8.4	571	12	US-11-128-061-6866	Sequence 6866, Ap
	35	185	8.4	571	12	US-11-128-049-6866	Sequence 6866, Ap
	36	185	8.4	651	12	US-11-128-061-3224	Sequence 3224, Ap
	37	185	8.4	651	12	US-11-128-049-3224	Sequence 3224, Ap
	38	182.2	8.3	520	12	US-11-128-061-5823	Sequence 5823, Ap
	39	182.2	8.3	520	12	US-11-128-049-5823	Sequence 5823, Ap
	40	179.2	8.2	515	12	US-11-000-688-393	Sequence 393, App
	41	178.8	8.2	387	8	US-10-821-234-792	Sequence 792, App
	42	177.4	8.1	428	12	US-11-000-688-392	Sequence 392, App
c	43	174.4	8.0	541	6	US-09-925-065A-634066	Sequence 634066,
	44	173.2	7.9	576	12	US-11-136-527-58	Sequence 58, Appl
c	45	173.2	7.9	576	12	US-11-136-527-4154	Sequence 4154, Ap

ALIGNMENTS

RESULT 1
US-11-050-857-1056
; Sequence 1056, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1056
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-050-857-1056

Query Match 99.3%; Score 2177; DB 9; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy          76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy         136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy         196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
```

```
Db       301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360
Qy       376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
Db       361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420
Qy       436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
Db       421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480
Qy       496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
Db       481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540
Qy       556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
Db       541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600
Qy       616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
Db       601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660
Qy       676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
Db       661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720
Qy       736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
Db       721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780
Qy       796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
Db       781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840
Qy       856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
Db       841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900
Qy       916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
Db       901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960
Qy       976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
Db       961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020
Qy      1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
Db      1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080
Qy      1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
Db      1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140
Qy      1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
Db      1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200
Qy      1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
Db      1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260
Qy      1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
Db      1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320
Qy      1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
Db      1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380
Qy      1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
Db      1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440
Qy      1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
Db      1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500
Qy      1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
Db      1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560
Qy      1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
Db      1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620
Qy      1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
Db      1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680
Qy      1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
Db      1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740
Qy      1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
Db      1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800
Qy      1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
Db      1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860
Qy      1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
Db      1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920
Qy      1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
Db      1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980
Qy      1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
Db      1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040
Qy      2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
Db      2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100
Qy      2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
Db      2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160
Qy      2176 AAAGTTGTTGCATTCCC 2192
Db      2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 2
US-11-043-590-46
; Sequence 46, Application US/11043590
; Publication No. US20050277156A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use there
; FILE REFERENCE: 1847.1011
; CURRENT APPLICATION NUMBER: US/11/043,590
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 46
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-043-590-46

Query Match 99.3%; Score 2177; DB 12; Length 2201;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA  75
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA  60

Qy      76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy     136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy     196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy     256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy     316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy     376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy     436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy     496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy     556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy     616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy     676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy     736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy     796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy     856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy     916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy     976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy    1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy    1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy    1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy    1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy    1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy    1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy    1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy    1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy    1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy    1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy    1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy    1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy    1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy    1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy    1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy    1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy    1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy    2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy    2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160
```

```
Qy       2176 AAAGTTGTTGCATTCCC 2192
              |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 3
US-11-136-527-2416
; Sequence 2416, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2416
; LENGTH: 2764
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2416

Query Match 75.0%; Score 1644.4; DB 12; Length 2764;
Best Local Similarity 81.3%; Pred. No. 0;
Matches 1703; Conservative 156; Mismatches 234; Indels 1; Gaps 1;

```
Qy         26 GTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAA 85
              |||  |   |||| | || |||  || |  | |  | | ||||||||||||||||| |||||
Db          3 GTCTCTACTCTTTTGTGAGGACTGCAGCCAACACCGC-TGACAATGCAGATCTTTGTGAA 61

Qy         86 GACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGT 145
               ||  | |||||||||||||||||||| ||||| |||||||||||||||||| || |||||
Db         62 AACCTTAACTGGTAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATTGAAAATGT 121

Qy        146 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGC 205
              |||||||||||||| || |||||  |||||||| |||||||||||||||||||||||||||||
Db        122 CAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGC 181

Qy        206 TGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 265
              || || |||||||||||||| |||||||||| |||||||||||||||||| ||||||||
Db        182 AGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAGGAGTCCAC 241

Qy        266 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 325
              |:|||||||||| || || |||||||| || ||||| ||||| |||||||| || || ||
Db        242 CYTGCACCTGGTCCTKCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGACCCTGACAGG 301

Qy        326 CAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGAT 385
              |||||||||||||||| ||||| |||||||||||||||| || || ||||| ||||||||
Db        302 CAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATTGAAAATGTCAAGGCAAAGAT 361

Qy        386 CCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              |||||||||||| ||||| || |||||||||||||||| |||||||||| || ||||||||
Db        362 CCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCT 421

Qy        446 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 505
              ||||||||| |||||||||| |||||||||||||||| ||||| ||| ||||||||||
Db        422 GGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAGGAGTCCACCTTGCACCTGGT 481

Qy        506 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 565
              || || |||||||| || |||||||||| |||||||||||||||| || ||||||||||
Db        482 CCTTCGCCTCAGAGGGGGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCAC 541

Qy        566 CCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGA 625
              ||| ||||| |||||||:||||||||| || |||||||||||||||||| || :||:|
Db        542 CCTGGAGGTCGAGCCCARTGACACCATTGAAAATGTCAAGGCAAAGATCCAGGACWAGKA 601

Qy        626 AGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 685
              ||||| || ||||| |||||:||| |||||||||| || || |||:|||||||||| ||
```

```
Db        602 GGGCATCCCCCCTGACCAGCASAGGCTGATCTTTGCAGGCAAGCAGSTGGAAGATGGCCG 661

Qy        686 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAG 745
              |||||||| |||||||||||||||| |||||||| |||||||||| || || |||||
Db        662 CACCCTGTCAGACTACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCCTTCGCCTCAG 721

Qy        746 AGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGA 805
              ||| || ||||| ||||| |||||||| || || |||||||:|||||||| ||||||||
Db        722 AGGGGGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGAMCATCACCCTGGAGGTCGA 781

Qy        806 GCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCC 865
              |||||||||||| || || || ||||| |||||||||| |:|||||| ||||| || ||
Db        782 GCCCAGTGACACCATTGAAAATGTCAAGGCAAAGATCCAGGWCAAGGAGGGCATCCCCCC 841

Qy        866 TGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGA 925
              |||||||||||| |||||||||| ||:|||||:||:|:||:|:| |||||||||| ||
Db        842 TGACCAGCAGAGGCTGATCTTTGCAGGMAAGCARCTKGWAGRTKGCCGCACCCTGTCAGA 901

Qy        926 CTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 985
              |||||||||||||||| ||||| |||||||||||||| || || |||||||| || |||||
Db        902 CTACAACATCCAGAAGGAGTCCACCCTGCACCTGGTCCTGCGCCTCAGAGGGGGCATGCA 961

Qy        986 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACAC 1045
              |||||| |||||||||||||| || |||||||||||||| || || ||||| ||||||||
Db        962 GATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACAC 1021

Qy       1046 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1105
              |||||| |||||||||||||||||| |||||||| |||||||| ||||||||||||||
Db       1022 CATTGAAAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAG 1081

Qy       1106 GTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCA 1165
              | |||||||||| || || ||:|||||::||| || ||||||:|: ||||:|||||:||
Db       1082 GCTGATCTTTGCAGGCAAGCARCTGGAARWTGGCCGCACCCTKTYAGACTWCAACATYCA 1141

Qy       1166 GAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 1225
              ||| |||::|||| |::|||||: || :| ||:|:||| || ||||:||:||| |:|||
Db       1142 GAAGGAGKYCACCCTKSACCTGGKCCTTYGVCTYARAGGGGGCATGCMGAWCTTTGKGAA 1201

Qy       1226 GACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGT 1285
              :|:|||::| || ||:|:|||:|: || |:||| ||||| |::|:|||||:||| ||:|:
Db       1202 RAMCCTKRCAGGCAARAMCATYAMCCTKGMGGTCGAGCCCARKGRCACCAWTGAAAAWGK 1261

Qy       1286 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 1345
              :|||||||:||||| |||||||: |||||:|| |||:| ||:||:|:|: |:|||||:|
Db       1262 YAAGGCAAARATCCAGGACAAGGRGGGCATYCCCCCTKACCARCARARGCTKATCTTTKC 1321

Qy       1346 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1405
              :| ||:||:||:|||:|||| ||:|||||||:| :|||:|||||||:|:|| |||||:||
Db       1322 ARGCAARCARCTKGAARATGGCCGMACCCTGKCARACTWCAACATCMARAAGGAGTCMAC 1381

Qy       1406 CCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1465
              |:|||||||||| || :| ||:|:||| || |||||||||| ::|||:|||||::| :|
Db       1382 CYTGCACCTGGTCCTTYGSCTYARAGGGGGCATGCAGATCTTTKKGAARACCCTKRCASG 1441

Qy       1466 TAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGAT 1525
              |||||||||||| || || ||||| ||:|||||||||: |||||:|||||::|||
Db       1442 CAAGACCATCACCCTKGAGGTCGAGCCCAGKGACACCATTGRAAATGTSAAGGCAWWGAT 1501

Qy       1526 CCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 1585
              :|| |||||||| ||||::|| |||||||||||||||| |||||||||| :| || |||||
Db       1502 YCAKGACAAGGAGGGCAWHCCCCCTGACCAGCAGAGGCTGATCTTTGCAKGCAAGCAGCT 1561

Qy       1586 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
              :|:||||:| ||||||||||: ||||:|||||:|||:|:|||::||:||||:||||:|:
Db       1562 KGRAGATKGSCGCACCCTGTYAGACTWCAACATYCAGRARGAGKSCAMCCTGSACCTSGK 1621

Qy       1646 GCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1705
              || :| || |:||| || |||||||||| |||||||||||||||| :| :|||||||||
Db       1622 YCTTYGSCTCARAGGGGGCATGCAGATCTTTGTGAAGACCCTGACASGCWAGACCATCAC 1681

Qy       1706 TCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1765
              || ||:|| ||||| |||||||||||||| |||||||||||||||||||| ||||||||
Db       1682 CCTGGARGTYGAGCCCAGTGACACCATTGAAAATGTCAAGGCAAAGATCCAGGACAAGGA 1741
```

```
Qy       1766 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 1825
Db       1742 GGGCATCCYCCCTGACCAGCAGAGGCTGATCTTTGCMGGCAAGCAGCTKGAASATGRCCG 1801

Qy       1826 CACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAG 1885
Db       1802 CACCCTGTCAGACTACAACATMCARAAGGAGTCCACCCTGCACCTGGTYCTGYGYCTMAK 1861

Qy       1886 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 1945
Db       1862 AGGGKGCATGCRGATMTTTGTGAAGACCCTGACAGGCWAGACCATCACCCTSGRGGKSGA 1921

Qy       1946 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 2005
Db       1922 GCCCAGTGACACCATTGAAAAYGTMAAGGCAVASRTCCAGGACAAGGWGGGSATCCCCCC 1981

Qy       2006 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 2065
Db       1982 TGAMCWSSARAGSSTGATCTTTGCAGGCAAGCAGYTGGAAGATGSCCGCACCCTGTCAGA 2041

Qy       2066 CTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
Db       2042 YKRCAMSAKSCRGMASGAGKCCACCCTGCAMYTGGWCMKKCGCYTCAGAGGGGG 2095
```

RESULT 4
US-11-102-026A-8
; Sequence 8, Application US/11102026A
; Publication No. US20060021087A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Baum, James A
; APPLICANT: Gilbertson, Larry A
; APPLICANT: Kovalic, David K
; APPLICANT: LaRosa, Thomas J
; APPLICANT: Lu, Maolong
; APPLICANT: Munyikwa, Tichifa R. I.
; APPLICANT: Roberts, James K
; APPLICANT: Wu, Wei
; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: Compositions and Methods for Control of Insect Infestations in Pl
; FILE REFERENCE: 38-21(53596)
; CURRENT APPLICATION NUMBER: US/11/102,026A
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 60560842
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 60565632
; PRIOR FILING DATE: 2004-04-27
; PRIOR APPLICATION NUMBER: 60579062
; PRIOR FILING DATE: 2004-06-11
; PRIOR APPLICATION NUMBER: 60603421
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60617261
; PRIOR FILING DATE: 2004-10-11
; PRIOR APPLICATION NUMBER: 60669241
; PRIOR FILING DATE: 2005-04-07
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 8
; LENGTH: 2346
; TYPE: DNA
; ORGANISM: Lygus hesperus
US-11-102-026A-8

Query Match 66.0%; Score 1447.8; DB 12; Length 2346;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1680; Conservative 0; Mismatches 375; Indels 1; Gaps 1;

```
Qy         96 GGTAAGACCATCACCCTCGAGGTTGAGCCCAGT-GACACCATCGAGAATGTCAAGGCAAA 154
Db        124 GGGAAGACCATCACCCTCGAGGTCGAGCCTTCTCGATACCATTGAAAACGTGAAGGCGAA 183

Qy        155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
Db        184 AATTCAGGATAAAGAAGGCATCCCCCCAGATCAGCAGAGGTTGATCTTTGCCGGCAAGCA 243

Qy        215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
Db        244 GTTGGAAGACGGACGTACTTTGTCTGACTACAACATCCAAAAAGAATCCACTCTCCACCT 303

Qy        275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
Db        304 GGTCTTGAGATTGAGAGGTGGCATGCAGATCTTCGTGAAGACCCTCACAGGAAAGACCAT 363

Qy        335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
Db        364 CACTCTTGAGGTCGAGCCTTCTGACTCCATCGAAAACGTCAAGGCTAAAATTCAAGACAA 423

Qy        395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGG 454
Db        424 GGAAGGTATTCCTCCAGATCAGCAGAGATTGATCTTCGCCGGCAAACAACTCGAAGATGG 483

Qy        455 GCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCT 514
Db        484 CCGTACCCTCTCTGACTACAATATTCAAAAAGAGTCCACCCTTCACTTGGTGTTGAGATT 543

Qy        515 CAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGT 574
Db        544 GCGTGGAGGTATGCAAATCTTTGTCAAAACATTGACTGGAAAGACCATCACCCTTGAAGT 603

Qy        575 GGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCC 634
Db        604 CGAACCCTCCGACACCATCGAAAATGTCAAGGCCAAGATCCAGGACAAGGAAGGCATCCC 663

Qy        635 TCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTC 694
Db        664 CCCAGATCAGCAGAGGTTGATTTTCGCTGGCAAACAACTTGAAGACGGACGTACCCTCTC 723

Qy        695 TGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGAT 754
Db        724 GGACTACAACATCCAGAAGGAGTCGACCCTCCATCTTGTCCTCCGTCTGCGTGGTGGTAT 783

Qy        755 GCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGA 814
Db        784 GCAGATTTTTGTCAAAACTCTGACTGGCAAGACAATCACCCTTGAAGTAGAGCCCTCTGA 843

Qy        815 CACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCA 874
Db        844 CACCATCGAAAATGTCAAGGCGAAAATCCAGGACAAAGAAGGCATCCCCCCAGATCAGCA 903

Qy        875 GAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 934
Db        904 GAGGTTGATCTTCGCCGGTAAGCAGCTTGAAGACGGCCGTACCCTCTCGGACTACAACAT 963

Qy        935 CCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGT 994
Db        964 CCAGAAGGAGTCCACCCTTCATCTTGTCCTCCGTCTGCGTGGTGGTATGCAGATTTTCGT 1023

Qy        995 GAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAA 1054
Db       1024 GAAGACCTTGACTGGCAAGACCATCACTCTTGAGGTCGAGCCCTCTGACACCATCGAAAA 1083

Qy       1055 TGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTT 1114
Db       1084 CGTCAAGGCCAAGATCCAGGACAAGGAAGGTATCCCCCCAGATCAGCAGAGGTTGATCTT 1143

Qy       1115 TGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTC 1174
Db       1144 CGCTGGCAAGCAGCTCGAGGATGGTCGTACCCTCTCGGACTACAACATCCAGAAGGAGTC 1203

Qy       1175 CACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1234
Db       1204 CACCCTTCATCTTGTCCTCCGTCTGCGTGGTGGTATGCAGATTTTCGTGAAGACCTTGAC 1263

Qy       1235 TGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1294
```

```
Db       1264 TGGCAAGACCATCACTCTTGAGGTCGAGCCCTCTGACACCATTGAAAACGTCAAGGCCAA 1323

Qy       1295 GATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 1354

Db       1324 GATCCAGGACAAGGAAGGTATCCCCCCAGATCAGCAGAGGTTGATCTTCGCCGGTAAGCA 1383

Qy       1355 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1414

Db       1384 GCTTGAAGACGGCCGTACTCTCTCTGATTACAACATCCAGAAGGAGTCGACCCTCCACCT 1443

Qy       1415 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1474

Db       1444 TGTCCTCCGTCTGCGTGGTGGTATGCAGATTTTCGTGAAGACCTTGACTGGCAAGACCAT 1503

Qy       1475 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1534

Db       1504 CACTCTTGAGGTCGAGCCCTCTGACACCATTGAAAACGTCAAGGCCAAGATCCAGGATAA 1563

Qy       1535 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1594

Db       1564 GGAAGGCATCCCCCCAGATCAGCAGAGGTTGATCTTCGCCGGTAAGCAGCTTGAGGATGG 1623

Qy       1595 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654

Db       1624 ACGTACCCTGTCAGACTACAACATCCAAAAGGAGTCCACCCTGCACTTGGTGTTGAGATT 1683

Qy       1655 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 1714

Db       1684 GCGTGGTGGTATGCAGATCTTCGTCAAGACCTTGACTGGCAAGACGATCACTTTGGAAGT 1743

Qy       1715 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1774

Db       1744 CGAGCCCTCTGACACCATTGAGAATGTCAAAGCCAAAATCCAAGATAAGGAAGGCATCCC 1803

Qy       1775 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1834

Db       1804 CCCAGATCAGCAGAGGTTGATCTTCGCTGGTAAGCAGCTTGAAGACGGCCGCACTCTTTC 1863

Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894

Db       1864 GGATTACAACATCCAGAAGGAGTCGACCCTCCACCTTGTCCTTCGTCTGCGTGGTGGTAT 1923

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954

Db       1924 GCAGATCTTCGTCAAGACGTTGACAGGCAAGACCATCACCCTTGAAGTCGAGCCCTCTGA 1983

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014

Db       1984 CACCATCGAAAACGTCAAGGCTAAGATCCAGGACAAGGAAGGTATCCCCCCAGATCAGCA 2043

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074

Db       2044 AAGATTGATCTTCGCCGGCAAACAGCTCGAAGATGGCCGTACCCTCTCAGACTACAACAT 2103

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 2134

Db       2104 TCAAAAGGAGTCAACTCTTCATCTCGTTCTGAGGCTCCGTGGCGGTCGTTATTGATCACA 2163

Qy       2135 TTTTAAGGTTTCAACA 2150

Db       2164 ATTCCAAACTTAAAAA 2179


RESULT 5
US-11-117-187-211
; Sequence 211, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
;  APPLICANT: PREUSS, DAPHNE
;  APPLICANT:  COPENHAVER, GREGORY
;  TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;  FILE REFERENCE: ARCD:309US
;  CURRENT APPLICATION NUMBER: US/11/117,187
;  CURRENT FILING DATE:  2005-04-28
;  PRIOR APPLICATION NUMBER: US/09/531,120
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/125,219
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
;   LENGTH: 1082144
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-11-117-187-211

  Query Match             41.5%;  Score 909.4;  DB 12;  Length 1082144;
  Best Local Similarity   79.2%;  Pred. No. 5.2e-253;
  Matches 1092;  Conservative    0;  Mismatches  286;  Indels    1;  Gaps    1;

Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583

Db     747402 GATGCAGATCTTTGTTAAGACTCTCACCGGAAAGACAATCACCCTCGAGGTGGAAAGCTC 747461

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643

Db     747462 CGACACCATCGACAACGTTAAGGCCAAGATCCAGGATAAGGAGGGCATTCCTCCGGATCA 747521

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703

Db     747522 GCAGAGGCTTATTTTCGCCGGCAAGCAGCTAGAGGATGGCCGTACGTTGGCTGATTACAA 747581

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763

Db     747582 TATCCAGAAGGAATCCACCCTCCACTTGGTCCTCAGGCTCCGTGGTGGTATGCAGATTTT 747641

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823

Db     747642 CGTTAAAACCCTAACGGGAAAGACGATTACTCTTGAGGTGGAGAGTTCTGACACCATCGA 747701

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883

Db     747702 CAACGTCAAGGCCAAGATCCAAGACAAAGAGGGTATTCCTCCGGACCAGCAGAGGCTGAT 747761

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943

Db     747762 CTTCGCCGGAAAGCAGTTGGAGGATGGCAGAACTCTTGCTGACTACAATATCCAGAAGGA 747821

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003

Db     747822 GTCCACCCTTCATCTTGTTCTCAGGCTCCGTGGTGGTATGCAGATTTTCGTTAAGACGTT 747881

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063

Db     747882 GACTGGGAAAACTATCACTTTGGAGGTGGAGAGTTCTGACACCATTGATAACGTGAAAGC 747941

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123

Db     747942 CAAGATCCAAGACAAAGAGGGTATTCCTCCGGACCAGCAGAGATTGATCTTCGCCGGAAA 748001

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183

Db     748002 ACAACTTGAAGATGGCAGAACTTTGGCCGACTACAACATTCAGAAGGAGTCCACACTCCA 748061

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243

Db     748062 CTTGGTCTTGCGTCTGCGTGGAGGTATGCAGATCTTCGTGAAGACTCTCACCGGAAAGAC 748121

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303

Db     748122 CATCACTTTGGAGGTGGAGAGTTCTGACACCATTGATAACGTGAAAGCCAAGATCCAGGA 748181

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363

Db     748182 CAAAGAGGGTATCCCACCGGACCAGCAGAGATTGATCTTCGCCGGAAAGCAACTTGAAGA 748241
```

```
Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423
             |||| | ||  || |||||||||||||| ||||| |||||||| || ||| ||||  | ||
Db     748242 TGGAAGAACTTTGGCTGACTACAACATTCAGAAGGAGTCCACACTTCACTTGGTCTTGCG 748301

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1483
             |||  | || || |||||||||||||||||||| || || || ||||| |||||| | ||
Db     748302 TCTGCGTGGAGGTATGCAGATCTTCGTGAAGACTCTCACCGGAAAGACTATCACTTTGGA 748361

Qy       1484 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1543
              || |||      |||||||||| || || ||||| |||||||| || |||||||| ||
Db     748362 GGTAGAGAGCTCTGACACCATTGACAACGTGAAGGCCAAGATCCAGGATAAGGAAGGAAT 748421

Qy       1544 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1603
             |||||| |||||||||||||||||||||| || |||||  |||| ||||| || ||  |
Db     748422 CCCTCCGGACCAGCAGAGGTTGATCTTTGCCGGAAAACAATTGGAGGATGGTCGTACTTT 748481

Qy       1604 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1663
             | | || |||||||||||||| ||||| ||||| ||| ||||| | |||||  | || ||
Db     748482 GGCGGATTACAACATCCAGAAGGAGTCGACCCTTCACTTGGTGTTGCGTCTGCGTGGAGG 748541

Qy       1664 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAG 1723
              |||||||||||||| |||||  |||| || |||||||||| || |||||||||
Db     748542 TATGCAGATCTTCGTCAAGACTTTGACCGGAAAGACCATCACCCTTGAAGTGGAAAGCTC 748601

Qy       1724 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 1783
              |||||||||| || |||||||| |||||||| ||||||| ||| || ||||| |||||
Db     748602 CGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGG-AGGTATTCCTCCGGACCA 748660

Qy       1784 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1843
             |||| |  | ||||| ||||| || ||||| || |||||||| ||  || | ||||||||
Db     748661 GCAGCGTCTCATCTTCGCTGGAAAGCAGCTTGAGGATGGACGTACTTTGGCCGACTACAA 748720

Qy       1844 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 1902
             ||||||||| ||||| || || ||| |||| || |||||  | |||||  |  | ||||
Db     748721 CATCCAGAAGGAGTCTACTCTTCACTTGGTCCTGCGTCTTCGTGGTGGTTTCTAAATCT 748779


RESULT 6
US-11-136-527-6512
; Sequence 6512, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6512
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-6512

  Query Match           37.0%;  Score 810;  DB 12;  Length 1400;
  Best Local Similarity   76.9%;  Pred. No. 2.5e-225;
  Matches  800;  Conservative  129;  Mismatches  111;  Indels    0;  Gaps    0;

Qy        705 ATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTC 764
              |||:|:|| |||||:|||:|||||||||| || :| ||:|:||| || ||||| |||||
Db          1 ATCMARAAGGAGTCMACCYTGCACCTGGTCCTTYGSCTYARAGGGGGCATGCAGATCTTT 60

Qy        765 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAG 824
              ::|||:|| || :| :|||||||||||||||:|||||||||||||:||||| || |:
Db         61 KKGAARACCCTKRCASGCAAGACCATCACCCTKGAGGTCGAGCCCAGKGACACCATTGRA 120

Qy        825 AACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 884
              || ||:|| |||::|||:|| |||||||| ||||::|| |||||||||||||| |||||
Db        121 AATGTSAAGGCAWWGATYCAKGACAAGGAGGGCAWHCCCCCTGACCAGCAGAGGCTGATC 180

Qy        885 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 944
              ||||| :| |||||||||:|:||||:|:||||||||||: ||||:||||||:|||:|:|||
Db        181 TTTGCAKGCAAGCAGCTKGRAGATKGSCGCACCCTGTYAGACTWCAACATYCAGRARGAG 240

Qy        945 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1004
              :: |:||||:||||:|: || :| ||||:||| || |||||||||| ||||||||||
Db        241 KSCAMCCTGSACCTSGKYCTTYGSCTCARAGGGGGCATGCAGATCTTTGTGAAGACCCTG 300

Qy       1005 ACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1064
              || :| :|||||||||||||| ||:|| ||||| |||||||||||||| ||||||||||
Db        301 ACASGCWAGACCATCACCCTGGARGTYGAGCCCAGTGACACCATTGAAAATGTCAAGGCA 360

Qy       1065 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAA 1124
              |||||||| |||||||| |||||||: |||||||||||||| ||||||||||:|| ||
Db        361 AAGATCCAGGACAAGGAGGGCATCCYCCCTGACCAGCAGAGGCTGATCTTTGCMGGCAAG 420

Qy       1125 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 1184
              |||||:|||:|||: || |||||||| ||||||||||||:||:|| |||||||| |||||
Db        421 CAGCTKGAAASATGRCCGCACCCTGTCAGACTACAACATMCARAAGGAGTCCACCCTGCAC 480

Qy       1185 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1244
              ||||| || :|:||:|:||| :| |||:|:||:|| |||||||||||||| || :|||||
Db        481 CTGGTYCTGYGYCTMAKAGGGKGCATGCRGATMTTTGTGAAGACCCTGACAGGCWAGACC 540

Qy       1245 ATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 1304
              ||||| ||:|:||::||||| |||||||||||||| ||:||:||||||:|::|||| |||
Db        541 ATCACCCTSGRGGKSGAGCCCAGTGACACCATTGAAAAYGTMAAGGCAVASRTCCAGGAC 600

Qy       1305 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 1364
              ||||: ||:||||| ||||| |:::|:||: |||||||||| || || |||:|||||||
Db        601 AAGGWGGGSATCCCCCCTGAMCWSSARAGSSTGATCTTTGCAGGCAAGCAGYTGGAAGAT 660

Qy       1365 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGT 1424
              |: |||||||||| ||:::||::|::|:|:| |||:||||||||::|||: :: ||
Db        661 GSCCGCACCCTGTCAGAYKRCAMSAKSCRGMASGAGKCCACCCTGCAMYTGGWCMKKCGC 720

Qy       1425 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 1484
              :| ||||| || |||||||||| |||||||||:|||| || |||:||||||| :| ||
Db        721 YTCAGAGGGGGCATGCAGATCTTTGTGAAGACCYTGACAGGCAAGMCCATCACCYTGGAG 780

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
              || ::||| |||||||||:||:::||:|:::|||||||||||| |:|||||| ||||||
Db        781 GTCRRGCCCAGTGACACCMTTSRRAAYGWSRAGGCAAAGATCCAGGRCAAGGAGGGCATC 840

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
              || |||||||||||:|||:|||||||||| || || |||:|||||||||: |:|||||||
Db        841 CCCCCTGACCAGCARAGGYTGATCTTTGCAGGCAAGCAGYTGGAAGATGSCCSCACCCTG 900

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1664
              || |:||||::||||:||:| |||||||||:|||||||:| :| || |: |||:: ::
Db        901 TCAGRCTACRRCATCSAGRAGGAGTCCACCYTGCACCTGRTCYTKCGCCYCAGASSRSRM 960

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
              ::|||||||:| |||||||||||||| || |||:|||||||:| || || ||||| |||
Db        961 RKGCAGATCWTTGTGAAGACCCTGACAGGCAAGRCCATCACTYTGGAKGTCGAGCCCAGT 1020

Qy       1725 GACACCATTGAGAATGTCAA 1744
              |  | |:   |||| |||||
Db       1021 GTTAACMCCAAGAAGGTCAA 1040


RESULT 7
US-11-136-527-1813
; Sequence 1813, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
```

```
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1813
;   LENGTH: 1223
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (825)..(825)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (930)..(930)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1065)..(1065)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1163)..(1163)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-1813

  Query Match             32.8%;  Score 718.4;  DB 12;  Length 1223;
  Best Local Similarity   75.8%;  Pred. No. 1e-198;
  Matches  726;  Conservative  113;  Mismatches  119;  Indels    0;  Gaps    0;

Qy         36 TTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACT 95
              |||||  |  |||||  |  |   ||  | | ||||||  ||:|||||||||| |||||
Db         66 TTGTTCCTTCATCGCATTCGTTAACAGGTCAAAATGCAAATYTTCGTGAAGACCCTGACC 125

Qy         96 GGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 155
              || |||||||||||||| ||||| |||||||||||||||||||||||| || ||||| |||
Db        126 GGCAAGACCATCACCCTAGAGGTGGAGCCCAGTGACACCATCGAGAACGTGAAGGCCAAG 185

Qy        156 ATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAG 215
              ||||| ||||| || |||||||| |||||||||||||||| |||||||| || || |||
Db        186 ATCCAGGATAAAGAGGGCATCCCCCCTGACCAGCAGAGGCTCATCTTTGCCGGCAAGCAG 245

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              |||||||||| |||||||| |:|||||||||||||||||||||||| |||||||:|||
Db        246 CTGGAAGATGGCCGCACCCTCTYTGACTACAACATCCAGAAAGAGTCAACCCTGCAMCTG 305

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              || ||||| || || || || ||||| |||||||||||||| || || ||||||||||||
Db        306 GTCCTCCGCCTKAGGGGSGGCATGCAGATCTTCGTGAAGACCCTGACCGGCAAGACCATC 365

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||| |||||||||||||||||||||||||||||| || || |||||||||| ||
Db        366 ACCCTGGAGGTGGAGCCCAGTGACACCATCGAGAACGTGAAGGCCAAGATCCAGGATAAA 425

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 455
              || ||||| || |||||||||||||||| | |||||||||| |:|||||:|||||||||
Db        426 GAGGGCATCCCCCCTGACCAGCAGAGGCTCATCTTTGCCGGCARGCAGCTKGAAGATGGC 485

Qy        456 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 515
              ||:||||| ||||| |||||||||||||||||||| |||||||||| || ||||| :|
Db        486 CGMACCCTCTCTGAWTACAACATCCAGAAAGAGTCAACCCTGCACCTTGTCCTCCGCYTG 545

Qy        516 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 575
              || || || ||||||||||||||:||:|||||||:|| ||||||||||||:| ||||||
Db        546 AGGGGCGGCATGCAGATCTTCGTRAARACCCTGACYGGCAAGACCATCACCYTRGAGGTG 605

Qy        576 GAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCT 635
              :|||||||::|:|||||||||: || |:|:| |||::|:| |:|:: |: |:||| ::
Db        606 KAGCCCAGTSMCMCCATCGAGAMCGTRARGSCCAAGMWCSAGGHTMRWGMSGSCATCYSC 665

Qy        636 CCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCT 695
              |:||| ||||||||::|:|||||||||||| || |||:|||||||||: || ||||| |||
Db        666 CYTGACCAGCAGAGSYTSATCTTTGCCGGCAAGCAGYTGGAAGATGSCCGMACCCTYTCT 725

Qy        696 GACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATG 755
              |::||||:||||:||||:||||| |||:::::|::|||: :|:|: :| :|:||:|| :|:
Db        726 GMYTACAMCATCMAGAARGAGTCAACCYWRMASBTGGWSYTYCRCYTGMGRGGYGGCMTS 785

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
              :| |::|::|||:|::| :: :: :|||:|:::|||::: : ||||| ||||||||||||
Db        786 YAGAWMTWBGTGRAKMCCBYGRSCKGCAVGMMYATCRSYNYGGAGGTGGAGCCCAGTGAC 845

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
              :| ||:|||||:|| || :| |||||:|: :| ||:|| ||:|| || ||||:|||||||
Db        846 MCCATSGAGAAYGTGAAGSCCAAGATMCWGRATAARGAGGGMATCCCCCCTGWCCAGCAG 905

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
              |||:|:||:|||:|||| |||||| |:||||:||: |||:|::|:|:||| ::|||||||
Db        906 AGGHTSATYTTTBCCGGCAAGCAGNTRGAAGWTGSCCGCRCYYTSTBTGATYMCAACATC 965

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 993
              ::|||:||||| |||:|:|:|:|||| |||||::| |: |||||:    |  |:|||:
Db        966 MRGAAVGAGTCRACCBTKCWCYTGGTCCTCCGYBTGAKGGGTGGVTATTAATTVTTCR 1023


RESULT 8
US-11-136-527-5909
; Sequence 5909, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William M
;  TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;  FILE REFERENCE: 031896-041000 (AM101086)
;  CURRENT APPLICATION NUMBER: US/11/136,527
;  CURRENT FILING DATE:  2005-05-25
;  PRIOR APPLICATION NUMBER: US 60/574,294
;  PRIOR FILING DATE: 2005-05-26
;  NUMBER OF SEQ ID NOS: 362830
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5909
;   LENGTH: 1223
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (825)..(825)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (930)..(930)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1065)..(1065)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1163)..(1163)
;   OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-5909

  Query Match             32.8%;  Score 718.4;  DB 12;  Length 1223;
  Best Local Similarity   75.8%;  Pred. No. 1e-198;
  Matches  726;  Conservative  113;  Mismatches  119;  Indels    0;  Gaps    0;

Qy         36 TTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACT 95
              |||||  |  |||||  |  |   ||  | | ||||||  ||:|||||||||| |||||
Db         66 TTGTTCCTTCATCGCATTCGTTAACAGGTCAAAATGCAAATYTTCGTGAAGACCCTGACC 125

Qy         96 GGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 155
              || |||||||||||||| ||||| |||||||||||||||||||||||| || ||||| |||
Db        126 GGCAAGACCATCACCCTAGAGGTGGAGCCCAGTGACACCATCGAGAACGTGAAGGCCAAG 185

Qy        156 ATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAG 215
              ||||| ||||| || |||||||| |||||||||||||||| |||||||| || || |||
```

```
Db        186 ATCCAGGATAAAGAGGGCATCCCCCCTGACCAGCAGAGGCTCATCTTTGCCGGCAAGCAG 245

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275

Db        246 CTGGAAGATGGCCGCACCCTCTYTGACTACAACATCCAGAAAGAGTCAACCCTGCAMCTG 305

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335

Db        306 GTCCTCCGCCTKAGGGGSGGCATGCAGATCTTCGTGAAGACCCTGACCGGCAAGACCATC 365

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395

Db        366 ACCCTGGAGGTGGAGCCCAGTGACACCATCGAGAACGTGAAGGCCAAGATCCAGGATAAA 425

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 455

Db        426 GAGGGCATCCCCCCTGACCAGCAGAGGCTCATCTTTGCCGGCARGCAGCTKGAAGATGGC 485

Qy        456 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 515

Db        486 CGMACCCTCTCTGAWTACAACATCCAGAAAGAGTCAACCCTGCACCTTGTCCTCCGCYTG 545

Qy        516 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 575

Db        546 AGGGGCGGCATGCAGATCTTCGTRAARACCCTGACYGGCAAGACCATCACCYTRGAGGTG 605

Qy        576 GAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCT 635

Db        606 KAGCCCAGTSMCMCCATCGAGAMCGTRARGSCCAAGMWCSAGGHTMRWGMSGSCATCYSC 665

Qy        636 CCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCT 695

Db        666 CYTGACCAGCAGAGSYTSATCTTTGCCGGCAAGCAGYTGGAAGATGSCCGMACCCTYTCT 725

Qy        696 GACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATG 755

Db        726 GMYTACAMCATCMAGAARGAGTCAACCYWRMASBTGGWSYTYCRCYTGMGRGGYGGCMTS 785

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815

Db        786 YAGAWMTWBGTGRAKMCCBYGRSCKGCAVGMMYATCRSYNYGGAGGTGGAGCCCAGTGAC 845

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875

Db        846 MCCATSGAGAAYGTGAAGSCCAAGATMCWGRATAARGAGGGMATCCCCCCTGWCCAGCAG 905

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935

Db        906 AGGHTSATYTTTBCCGGCAAGCAGNTRGAAGWTGSCCGCRCYYTSTBTGATYMCAACATC 965

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 993

Db        966 MRGAAVGAGTCRACCBTKCWCYTGGTCCTCCGYBTGAKGGGTGGVTATTAATTVTTCR 1023


RESULT 9
US-10-750-185-45440/c
; Sequence 45440, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 45440
;   LENGTH: 1499
;   TYPE: DNA
;   ORGANISM: Bovine  19866881311041
US-10-750-185-45440

  Query Match             25.8%;  Score 565;  DB 8;  Length 1499;
  Best Local Similarity   87.9%;  Pred. No. 4.6e-154;
  Matches  616;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy         64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123

Db       1402 TCAAAATGCAGATCTTCGTGAAAACCCTGACCGGCAAGACCATCACCCTGGAGGTGGAGC 1343

Qy        124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183

Db       1342 CCAGTGACACCATCGAGAACGTGAAGGCCAAGATCCAGGATAAGGAAGGCATTCCCCCTG 1283

Qy        184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243

Db       1282 ACCAGCAGAGGCTCATCTTTGCCGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGATT 1223

Qy        244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303

Db       1222 ACAACATCCAGAAAGAGTCGACCCTGGACCTGGTCCTCCGTCTGAGGGGTGGTATGCAGA 1163

Qy        304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363

Db       1162 TTTTCGTGAAGACCCTGACCGGCAAGACCATCACCCTGGAAGTGGAGCCCAGTGACACCA 1103

Qy        364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423

Db       1102 TCGAGAACGTGAAGGCCAAGATCCAGGATAAGGAGGGCATTCCCCCCGACCAGCAGAGGC 1043

Qy        424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483

Db       1042 TCATCTTTGCCGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGATTACAACATCCAGA 983

Qy        484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543

Db        982 AAGAGTCGACCCTGCACCTGGTCCTCCGTCTGAGGGGTGGTATGCAGATCTTCGTGAAGA 923

Qy        544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603

Db        922 CCCTGACCGGCAAGACCATCACCCTGGAGGTGGAGCCCAGTGACACCATCGAGAACGTGA 863

Qy        604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663

Db        862 AGGCCAAGATCCAGGATAAGGAAGGCATTCCCCCTGACCAGCAGAGGCTCATCTTTGCCG 803

Qy        664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723

Db        802 GCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGATTACAATATCCAGAAAGAGTCGACCC 743

Qy        724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTC 764

Db        742 TGCACCTGGTCCTTCGTCTGAGGGGTGGCTGTTAATTCTTC 702


RESULT 10
US-10-750-623-45440/c
; Sequence 45440, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
```

```
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 45440
;   LENGTH: 1499
;   TYPE: DNA
;   ORGANISM: Bovine  19866881311041
US-10-750-623-45440

  Query Match             25.8%;  Score 565;  DB 8;  Length 1499;
  Best Local Similarity   87.9%;  Pred. No. 4.6e-154;
  Matches  616;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy          64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
               | | ||||||||||||||||| || ||||| || ||||||||||||||| ||||| ||||
Db        1402 TCAAAATGCAGATCTTCGTGAAAACCCTGACCGGCAAGACCATCACCCTGGAGGTGGAGC 1343

Qy         124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
               |||||||||||||||||| || ||||| |||||||| |||||||||||||| || ||||
Db        1342 CCAGTGACACCATCGAGAACGTGAAGGCCAAGATCCAGGATAAGGAAGGCATTCCCCCTG 1283

Qy         184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
               |||||||||||| |||||||| || || |||||||||||||| ||||| || |||||
Db        1282 ACCAGCAGAGGCTCATCTTTGCCGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGATT 1223

Qy         244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
               |||||||||||||||||| |||||| ||||||| |||||||| || |||| |||||
Db        1222 ACAACATCCAGAAAGAGTCGACCCTGGACCTGGTCCTCCGTCTGAGGGGTGGTATGCAGA 1163

Qy         304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
               | |||||||||| || || |||||||||||||||| || ||||||||||||||||||||
Db        1162 TTTTCGTGAAGACCCTGACCGGCAAGACCATCACCCTGGAAGTGGAGCCCAGTGACACCA 1103

Qy         364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
               |||||||||| || || |||||||||| ||||| |||||||| || ||||||||||||
Db        1102 TCGAGAACGTGAAGGCCAAGATCCAGGATAAGGAGGGCATTCCCCCCGACCAGCAGAGGC 1043

Qy         424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
               | |||||||||| |||||||||||||||||| ||||| || ||||| ||||||||||||
Db        1042 TCATCTTTGCCGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGATTACAACATCCAGA 983

Qy         484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
               ||||||| |||||||||||||| |||||||| || ||||| ||||||||||||||||||
Db         982 AAGAGTCGACCCTGCACCTGGTCCTCCGTCTGAGGGGTGGTATGCAGATCTTCGTGAAGA 923

Qy         544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
               ||||||| || |||||||||||||| ||||||||||||||||||||||||||||| || |
Db         922 CCCTGACCGGCAAGACCATCACCCTGGAGGTGGAGCCCAGTGACACCATCGAGAACGTGA 863

Qy         604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
               |||| |||||||| |||||||||||||||||| ||||| ||||||||| | ||||||||||
Db         862 AGGCCAAGATCCAGGATAAGGAAGGCATTCCCCCTGACCAGCAGAGGCTCATCTTTGCCG 803

Qy         664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
               | || |||||||||||||| || || || ||||| ||||| |||||||||||||| |||
Db         802 GCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGATTACAATATCCAGAAAGAGTCGACCC 743

Qy         724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTC 764
               |||||||||| || ||||| || |||||     || |||||
Db         742 TGCACCTGGTCCTTCGTCTGAGGGGTGGCTGTTAATTCTTC 702


RESULT 11
US-11-122-329-68
; Sequence 68, Application US/11122329
; Publication No. US20060019272A1
; GENERAL INFORMATION:
;  APPLICANT: Geraci, Mark
;  APPLICANT:  Bull, Todd
;  APPLICANT:  Voelkel, Norbert
;  APPLICANT:  Coldren, Chris
;  TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
;  TITLE OF INVENTION:  Expression Analysis of Peripheral Blood Cells
;  FILE REFERENCE: 2848-54
;  CURRENT APPLICATION NUMBER: US/11/122,329
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/568,129
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 68
;   LENGTH: 3107
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-122-329-68

  Query Match             25.2%;  Score 552.2;  DB 12;  Length 3107;
  Best Local Similarity   86.7%;  Pred. No. 3.3e-150;
  Matches  608;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy          64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
               | | ||||||||||||||||| || || || || |||||||||||||| ||||| ||||
Db        2078 TCAAAATGCAGATCTTCGTGAAAACCCTTACCGGCAAGACCATCACCCTTGAGGTGGAGC 2137

Qy         124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
               |||||||||||||||| ||||| ||||| |||||||| |||||||||||||| || | |
Db        2138 CCAGTGACACCATCGAAAATGTGAAGGCCAAGATCCAGGATAAGGAAGGCATTCCCCCCG 2197

Qy         184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
               |||||||||||| |||||||| || || |||||||||||||| || || || |||||||
Db        2198 ACCAGCAGAGGCTCATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGTACTCTTTCTGACT 2257

Qy         244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
               |||||||||||| ||||| |||||||||||||| || ||||| |||||||| ||||| |
Db        2258 ACAACATCCAGAAGGAGTCGACCCTGCACCTGGTCCTGCGTCTGAGAGGTGGTATGCAGA 2317

Qy         304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
               |||||||||||| || || |||||||||||||||| || ||||||||||||||||||||
Db        2318 TCTTCGTGAAGACCCTGACCGGCAAGACCATCACCCTGGAAGTGGAGCCCAGTGACACCA 2377

Qy         364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
               |||| || || || || |||||||||| ||||| |||||||| ||||||||||||||||
Db        2378 TCGAAAATGTGAAGGCCAAGATCCAGGATAAAGAAGGCATCCCTCCCGACCAGCAGAGGC 2437

Qy         424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
               | |||||||| || |||||||||||||||||| ||||| || ||||||||||||||||||
Db        2438 TCATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAACATCCAGA 2497

Qy         484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
               | ||||| |||||||||||||| || ||||| |||||||| ||||||||||||||||||
Db        2498 AGGAGTCGACCCTGCACCTGGTCCTGCGTCTGAGAGGTGGTATGCAGATCTTCGTGAAGA 2557

Qy         544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
               ||||||| || |||||||||||||| || ||||| |||||||| |||||||||||||| |||
Db        2558 CCCTGACCGGCAAGACCATCACTCTGGAAGTGGAGCCCAGTGACACCATCGAAAATGTGA 2617

Qy         604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
               |||| |||||||||||||| |||||||| ||||| || ||||||||| | |||||||| |
Db        2618 AGGCCAAGATCCAAGATAAAGAAGGCATCCCTCCCGACCAGCAGAGGCTCATCTTTGCAG 2677

Qy         664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
               | || |||||||||||||| || || || |||||||||||||||||||| ||||| |||
Db        2678 GCAAGCAGCTGGAAGATGGCCGCACTCTTTCTGACTACAACATCCAGAAGGAGTCGACCC 2737

Qy         724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTC 764
               |||||||||| || || || || |||||     || |||||
Db        2738 TGCACCTGGTCCTGCGCCTGAGGGGTGGCTGTTAATTCTTC 2778


RESULT 12
US-10-927-641-1
```

```
; Sequence 1, Application US/10927641
; Publication No. US20050244968A1
; GENERAL INFORMATION:
;  APPLICANT: Perera, Ranjan
;  APPLICANT:  Rice, Stephen
;  APPLICANT:  Eagleton, Clare
;  APPLICANT:  Lasham, Annette
;  APPLICANT:  Wood, Marion
;  APPLICANT:  Visser, Elizabeth
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Expression
;  FILE REFERENCE: 11000.1036c4
;  CURRENT APPLICATION NUMBER: US/10/927,641
;  CURRENT FILING DATE:  2004-08-27
;  PRIOR APPLICATION NUMBER: PRIOR  APPLICATION NUMBER: US/10/137,036
;  PRIOR FILING DATE: 2002-04-30
;  PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
;  PRIOR FILING DATE: 2001-06-20
;  PRIOR APPLICATION NUMBER: U.S. No. 09/724,624
;  PRIOR FILING DATE: 2000-11-28
;  PRIOR APPLICATION NUMBER: U.S. No. 09/598,401
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: PCT/NZ00/00018
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: U.S. No. 60/146,591
;  PRIOR FILING DATE: 1999-07-30
;  PRIOR APPLICATION NUMBER: U.S. No. 09/276,599
;  PRIOR FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 143
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3083
;   TYPE: DNA
;   ORGANISM: Pinus radiata
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (1)...(2064)
;   FEATURE:
;   NAME/KEY: intron
;   LOCATION: (1196)...(2033)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2065)...(2751)
;   FEATURE:
;   NAME/KEY: 3'UTR
;   LOCATION: (2755)...(3083)
US-10-927-641-1

  Query Match            21.6%;  Score 474;  DB 8;  Length 3083;
  Best Local Similarity  77.2%;  Pred. No. 1.9e-127;
  Matches  576;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;

Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||| |  |  || |  || | |   |   ||    | || |  ||||||||||| |
Db       2018 TTGTGATGATGTGCAGGTATATTCGTTAGTCATATTTCAATTTCAAGATGCAGATCTTTG 2077

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              | |||||||| || |||||||||||||| |||||||| |||   |  |||||||| || |
Db       2078 TCAAGACTCTCACCGGTAAGACCATCACTCTCGAGGTCGAGAGCTCTGACACCATTGACA 2137

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              |||| || || |||||||| || |||||||| || || || ||||||||| | |||||||
Db       2138 ATGTTAAAGCTAAGATCCAGGACAAGGAAGGGATTCCCCCCGACCAGCAGCGTCTGATCT 2197

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              | || ||||| ||||| || || || || |||||  | || ||||||||||||||||| |
Db       2198 TCGCAGGAAAGCAGCTTGAGGACGGCCGAACCCTTGCCGATTACAACATCCAGAAAGAAT 2257

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              | ||||| ||||| || |||||| | || ||||| |||||||||| || || ||||| |
Db       2258 CTACCCTCCACCTTGTTCTCCGTTTGAGGGGTGGCATGCAAATCTTTGTAAAAACACTAA 2317

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              |||| ||||| || ||  | || || |||  |   |||||||| || |||||||| || |
Db       2318 CTGGAAAGACAATTACATTGGAAGTTGAGAGCTCGGACACCATTGACAACGTCAAGGCCA 2377

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||| ||||| |||||||||||||| | ||||| || || ||||
Db       2378 AGATCCAGGACAAGGAAGGAATTCCCCCTGACCAGCAGAGGCTTATCTTCGCTGGTAAGC 2437

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||| |||||  | ||| || |||| ||||| || || || || || ||||||||
Db       2438 AGCTGGAGGATGGCAGGACCTTGGCTGATTACAATATTCAAAAGGAATCGACCCTGCATT 2497

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||| ||||| ||||| || ||||| ||||| ||||| ||||| || |||||| ||||
Db       2498 TGGTGCTTCGTCTAAGAGGAGGCATGCAAATCTTTGTGAAAACCCTTACAGGTAAAACCA 2557

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              | || || || |||||   |   |||||||| || ||||| ||||| |||||||| || |
Db       2558 TTACTCTGGAAGTGGAAAGCTCGGACACCATTGACAATGTGAAGGCTAAGATCCAGGACA 2617

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              |||| || ||||| ||||| |||||||||||||||||||||| || ||||||||||||
Db       2618 AGGAGGGAATTCCACCTGACCAGCAGAGGTTGATCTTTGCCGGTAAGCAGCTGGAAGATG 2677

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||| ||  | || ||||| || || || ||| | |||||||| |||||||
Db       2678 GTCGTACTCTCGCCGATTACAATATTCAGAAGGAATCGACCCTTCACCTGGTGCTCCGTC 2737

Qy        742 TCAGAGGTGGGATGCAAATCTTCGTG 767
              || | |||||  |  |  |  |  |||
Db       2738 TCCGCGGTGGCTTTTAGGTTTGGGTG 2763


RESULT 13
US-10-927-641-143
; Sequence 143, Application US/10927641
; Publication No. US20050244968A1
; GENERAL INFORMATION:
;  APPLICANT: Perera, Ranjan
;  APPLICANT:  Rice, Stephen
;  APPLICANT:  Eagleton, Clare
;  APPLICANT:  Lasham, Annette
;  APPLICANT:  Wood, Marion
;  APPLICANT:  Visser, Elizabeth
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Expression
;  FILE REFERENCE: 11000.1036c4
;  CURRENT APPLICATION NUMBER: US/10/927,641
;  CURRENT FILING DATE:  2004-08-27
;  PRIOR APPLICATION NUMBER: PRIOR  APPLICATION NUMBER: US/10/137,036
;  PRIOR FILING DATE: 2002-04-30
;  PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
;  PRIOR FILING DATE: 2001-06-20
;  PRIOR APPLICATION NUMBER: U.S. No. 09/724,624
;  PRIOR FILING DATE: 2000-11-28
;  PRIOR APPLICATION NUMBER: U.S. No. 09/598,401
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: PCT/NZ00/00018
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: U.S. No. 60/146,591
;  PRIOR FILING DATE: 1999-07-30
;  PRIOR APPLICATION NUMBER: U.S. No. 09/276,599
;  PRIOR FILING DATE: 1999-03-25
;  NUMBER OF SEQ ID NOS: 143
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 143
;   LENGTH: 3084
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-10-927-641-143

  Query Match            21.6%;  Score 474;  DB 8;  Length 3084;
  Best Local Similarity  77.2%;  Pred. No. 1.9e-127;
  Matches  576;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;
```

```
Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||| |  |  || |  || | |    |    |    ||     | || |  ||||||||||| |
Db       2019 TTGTGATGATGTGCAGGTATATTCGTTAGTCATATTTCAATTTCAAGATGCAGATCTTTG 2078

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              | ||||||||| || ||||||||||||||| |||||||| |||   |  |||||||| || |
Db       2079 TCAAGACTCTCACCGGTAAGACCATCACTCTCGAGGTCGAGAGCTCTGACACCATTGACA 2138

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              |||| || || |||||||| || |||||||| || || || |||||||| | |||||||
Db       2139 ATGTTAAAGCTAAGATCCAGGACAAGGAAGGGATTCCCCCCGACCAGCAGCGTCTGATCT 2198

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              | || ||||| ||||| || || || || ||||| | || |||||||||||||||||| |
Db       2199 TCGCAGGAAAGCAGCTTGAGGACGGCCGAACCCTTGCCGATTACAACATCCAGAAAGAAT 2258

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              | ||||| ||||| || |||||| | || ||||| |||||||||| || || ||||| |
Db       2259 CTACCCTCCACCTTGTTCTCCGTTTGAGGGGTGGCATGCAAATCTTTGTAAAAACACTAA 2318

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              |||| ||||| || ||   | || || |||   |   |||||||| || |||||||| || |
Db       2319 CTGGAAAGACAATTACATTGGAAGTTGAGAGCTCGGACACCATTGACAACGTCAAGGCCA 2378

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||| ||||| |||||||||||||| | ||||| || || ||||
Db       2379 AGATCCAGGACAAGGAAGGAATTCCCCCTGACCAGCAGAGGCTTATCTTCGCTGGTAAGC 2438

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||| |||||  | ||| || |||| ||||| || || || || || ||||||||
Db       2439 AGCTGGAGGATGGCAGGACCTTGGCTGATTACAATATTCAAAAGGAATCGACCCTGCATT 2498

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||| ||||| ||||| || ||||| ||||| ||||| ||||| || ||||| ||||
Db       2499 TGGTGCTTCGTCTAAGAGGAGGCATGCAAATCTTTGTGAAAACCCTTACAGGTAAAACCA 2558

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              | || || || |||||   |   |||||||| || ||||| ||||| |||||||| || |
Db       2559 TTACTCTGGAAGTGGAAAGCTCGGACACCATTGACAATGTGAAGGCTAAGATCCAGGACA 2618

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              |||| || ||||| ||||| |||||||||||||||||||||||| || ||||||||||||
Db       2619 AGGAGGGAATTCCACCTGACCAGCAGAGGTTGATCTTTGCCGGTAAGCAGCTGGAAGATG 2678

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||| ||  | || ||||| || ||||| || || ||| | |||||||| |||||||
Db       2679 GTCGTACTCTCGCCGATTACAATATTCAGAAGGAATCGACCCTTCACCTGGTGCTCCGTC 2738

Qy        742 TCAGAGGTGGGATGCAAATCTTCGTG 767
              || | |||||  |  |  | |  |||
Db       2739 TCCGCGGTGGCTTTTAGGTTTGGGTG 2764


RESULT 14
US-11-117-187-211/c
; Sequence 211, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
;  APPLICANT: PREUSS, DAPHNE
;  APPLICANT:  COPENHAVER, GREGORY
;  TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;  FILE REFERENCE: ARCD:309US
;  CURRENT APPLICATION NUMBER: US/11/117,187
;  CURRENT FILING DATE:  2005-04-28
;  PRIOR APPLICATION NUMBER: US/09/531,120
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/125,219
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
;   LENGTH: 1082144
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-11-117-187-211

  Query Match             21.4%;  Score 469.4;  DB 12;  Length 1082144;
  Best Local Similarity   77.0%;  Pred. No. 5e-125;
  Matches  572;  Conservative    0;  Mismatches  171;  Indels    0;  Gaps    0;

Qy         21 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 80
              |||| |   |   ||||   |  ||| | ||  |  | | |   ||||||||||||
Db     617851 TTTGTGCGATCGAATTTGTCGATTAATCTCGGTTTTTCTGATTAACAGATGCAGATCTTC 617792

Qy         81 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 140
              || |||||||| || || ||||| ||||||||||||||| ||    |  ||||||||||||
Db     617791 GTTAAGACTCTCACCGGAAAGACTATCACCCTCGAGGTGGAAAGCTCTGACACCATCGAC 617732

Qy        141 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 200
              || || ||||| |||||||| |||||||||| || ||||| || |||||||||| |||
Db     617731 AACGTTAAGGCCAAGATCCAGGATAAGGAAGGTATTCCTCCGGATCAGCAGAGGCTTATC 617672

Qy        201 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 260
              || || ||||| ||| |||| |||||  |||||  || | || ||||| ||||||| ||
Db     617671 TTCGCCGGAAAGCAGTTGGAGGATGGCCGCACGTTGGCGGATTACAATATCCAGAAGGAA 617612

Qy        261 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 320
              |||||||| ||| |||| ||| | ||| | ||||| ||||| || ||||| || || ||
Db     617611 TCCACCCTCCACTTGGTTCTCAGGCTCCGTGGTGGTATGCAGATTTTCGTTAAAACCCTA 617552

Qy        321 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 380
              || || ||||| || || ||||||||||||   |  |||||||| || |||||||| ||
Db     617551 ACGGGAAAGACGATTACTCTTGAGGTGGAGAGCTCTGACACCATTGACAACGTCAAGGCC 617492

Qy        381 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 440
              |||||||| || ||||| || |||||||| ||||||||||||||||||| |||||||||
Db     617491 AAGATCCAAGATAAGGAGGGTATTCCTCCGGACCAGCAGAGGTTGATCTTCGCCGGAAAG 617432

Qy        441 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 500
              || || || || ||  | ||  || | || |||||||||||||||| |||||||| || ||
Db     617431 CAACTTGAGGACGGCAGAACTTTGGCGGATTACAACATCCAGAAGGAGTCTACGCTTCAT 617372

Qy        501 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 560
              ||||  | |||||  | || || |||||||||||||| |||||  |||| || ||||||
Db     617371 TTGGTCTTGCGTCTGCGTGGAGGTATGCAGATCTTCGTAAAGACTTTGACCGGAAAGACC 617312

Qy        561 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 620
              ||||| || || || ||| |   |||||||| || || || ||||| |||||||| ||
Db     617311 ATCACTCTTGAAGTTGAGAGCTCCGACACCATTGATAACGTGAAGGCTAAGATCCAGGAC 617252

Qy        621 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 680
              |||||||||||||||| || ||||| |  | ||||| || ||||| ||||| || |||
Db     617251 AAGGAAGGCATTCCTCCGGACCAGCAGCGTCTCATCTTCGCTGGAAAGCAGCTTGAGGAT 617192

Qy        681 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 740
              || |||||  || | ||||||||||||||||| ||||| ||  | ||| |||| ||||||
Db     617191 GGACGTACTTTGGCCGACTACAACATCCAGAAGGAGTCTACTCTTCACTTGGTCCTCCGT 617132

Qy        741 CTCAGAGGTGGGATGCAAATCTT 763
              ||| | |||||  |  ||| |||
Db     617131 CTCCGTGGTGGTTTCTAAACCTT 617109


RESULT 15
US-11-108-172-163
; Sequence 163, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Wang, Tongtong
```

; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-108-172-163

Query Match 19.8%; Score 433.4; DB 12; Length 460;
Best Local Similarity 99.8%; Pred. No. 5.3e-116;
Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         21 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 80
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 TGTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 85

Qy         81 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 145

Qy        141 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 205

Qy        201 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 265

Qy        261 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 325

Qy        321 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 385

Qy        381 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 445

Qy        441 CAGCTGGAAGATGGG 455
              |||||||||||||||
Db        446 CAGCTGGAAGATGGG 460
```

Search completed: March 6, 2006, 20:30:30
Job time : 662 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 19:39:04 ; Search time 370 Seconds
(without alignments)
10530.852 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1135.6	51.8	3840	2	US-08-462-092-1	Sequence 1, Appli
	2	1135.6	51.8	3840	3	US-08-746-822-1	Sequence 1, Appli
	3	1135.6	51.8	3840	3	US-09-094-350-1	Sequence 1, Appli
	4	1134	51.7	3840	9	5510474-1	Patent No. 5510474
	5	991.4	45.2	3184	3	US-09-516-277-1	Sequence 1, Appli
	6	988.8	45.1	4442	3	US-09-516-277-3	Sequence 3, Appli
	7	832.4	38.0	5174	3	US-09-866-153-8	Sequence 8, Appli
	8	832.4	38.0	5174	3	US-09-693-467A-8	Sequence 8, Appli
	9	832.4	38.0	5174	3	US-09-270-976-8	Sequence 8, Appli
	10	654.2	29.8	5512	3	US-09-866-153-5	Sequence 5, Appli
	11	654.2	29.8	5512	3	US-09-693-467A-5	Sequence 5, Appli
	12	654.2	29.8	5512	3	US-09-270-976-5	Sequence 5, Appli
c	13	644.2	29.4	795	2	US-08-592-126-140	Sequence 140, App
c	14	644.2	29.4	795	3	US-09-168-595-140	Sequence 140, App
	15	613.8	28.0	697	3	US-09-404-879A-100	Sequence 100, App
	16	613.8	28.0	697	3	US-09-338-933-100	Sequence 100, App
	17	613.8	28.0	697	3	US-09-215-681-100	Sequence 100, App
	18	613.8	28.0	697	3	US-09-216-003A-100	Sequence 100, App
	19	613.8	28.0	697	3	US-09-667-857-100	Sequence 100, App
	20	613.8	28.0	697	3	US-10-198-053-100	Sequence 100, App
	21	613.8	28.0	697	3	US-09-827-271-100	Sequence 100, App
c	22	592.8	27.0	636	3	US-09-404-879A-99	Sequence 99, Appl
c	23	592.8	27.0	636	3	US-09-338-933-99	Sequence 99, Appl
c	24	592.8	27.0	636	3	US-09-215-681-99	Sequence 99, Appl
c	25	592.8	27.0	636	3	US-09-216-003A-99	Sequence 99, Appl
c	26	592.8	27.0	636	3	US-09-667-857-99	Sequence 99, Appl
c	27	592.8	27.0	636	3	US-10-198-053-99	Sequence 99, Appl
c	28	592.8	27.0	636	3	US-09-827-271-99	Sequence 99, Appl
	29	568.2	25.9	585	3	US-09-404-879A-45	Sequence 45, Appl
	30	568.2	25.9	585	3	US-09-338-933-45	Sequence 45, Appl
	31	568.2	25.9	585	3	US-09-215-681-45	Sequence 45, Appl
	32	568.2	25.9	585	3	US-09-216-003A-45	Sequence 45, Appl
	33	568.2	25.9	585	3	US-09-667-857-45	Sequence 45, Appl
	34	568.2	25.9	585	3	US-10-198-053-45	Sequence 45, Appl
	35	568.2	25.9	585	3	US-09-827-271-45	Sequence 45, Appl
	36	564.6	25.8	1045	3	US-09-533-559-5864	Sequence 5864, Ap
	37	550.4	25.1	707	2	US-08-726-306A-22	Sequence 22, Appl
	38	529.6	24.2	547	3	US-09-370-838-171	Sequence 171, App
	39	529.6	24.2	547	3	US-09-854-133-171	Sequence 171, App
c	40	510.6	23.3	605	3	US-09-404-879A-95	Sequence 95, Appl
c	41	510.6	23.3	605	3	US-09-338-933-95	Sequence 95, Appl
c	42	510.6	23.3	605	3	US-09-215-681-95	Sequence 95, Appl
c	43	510.6	23.3	605	3	US-09-216-003A-95	Sequence 95, Appl
c	44	510.6	23.3	605	3	US-09-667-857-95	Sequence 95, Appl
c	45	510.6	23.3	605	3	US-10-198-053-95	Sequence 95, Appl

ALIGNMENTS

```
RESULT 1
US-08-462-092-1
; Sequence 1, Application US/08462092
; Patent No. 5614399
;  GENERAL INFORMATION:
;    APPLICANT:  Peter H. Quail
;    APPLICANT:  Alan H. Christensen
;    APPLICANT:  Howard P. Hershey
;    APPLICANT:  Robert A. Sharrock
;    APPLICANT:  Thomas D. Sullivan
;    TITLE OF INVENTION:  PLANT UBIQUITIN PROMOTER SYSTEM
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jeff Lloyd
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,092
;      FILING DATE:  5-JUNE-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lloyd, Jeff
;      REGISTRATION NUMBER:  35,589
;      REFERENCE/DOCKET NUMBER:  08/462,092
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3840 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1993..3591
US-08-462-092-1
```

Query Match 51.8%; Score 1135.6; DB 2; Length 3840;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 1315; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

```
Qy        64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
Db      1988 TGCAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATCACCCTCGAGGTGGAGT 2047

Qy       124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
Db      2048 CGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCAG 2107

Qy       184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
Db      2108 ACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGGCGCACGCTTGCTGACT 2167

Qy       244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
Db      2168 ACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTCAGGGGAGGCATGCAGA 2227

Qy       304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
Db      2228 TCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCCTCTGACACCA 2287

Qy       364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
Db      2288 TTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCAGCGGC 2347

Qy       424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
Db      2348 TCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGA 2407

Qy       484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
Db      2408 AGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGA 2467

Qy       544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
Db      2468 CCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAACGTCA 2527

Qy       604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
Db      2528 AGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGGCTCATCTTTGCTG 2587

Qy       664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
Db      2588 GAAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGAAGGAGAGCACCC 2647

Qy       724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 783
Db      2648 TCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCA 2707

Qy       784 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
Db      2708 AGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAATGTCAAGGCCAAGATCC 2767

Qy       844 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 903
Db      2768 AGGACAAGGAGGGCATCCCACCGGACCAGCAGCGTTTGATCTTCGCTGGCAAGCAGCTGG 2827

Qy       904 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 963
Db      2828 AGGATGGCCGCACCCTTGCGGATTACAACATCCAGAAGGAGAGCACCCTCCACCTGGTGC 2887

Qy       964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
Db      2888 TCCGTCTCAGGGGTGGTATGCAGATCTTTGTGAAGACACTCACTGGCAAGACAATCACCC 2947

Qy      1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083
Db      2948 TTGAGGTGGAGTCTTCGGATACCATTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGG 3007

Qy      1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143
Db      3008 GCATCCCACCCGACCAGCAGCGCCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCA 3067

Qy      1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203
Db      3068 CCCTGGCGGATTACAACATCCAGAAGGAGAGCACTCTCCACCTGGTGCTCCGCCTCAGGG 3127

Qy      1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263
Db      3128 GTGGCATGCAGATTTTTGTGAAGACATTGACTGGCAAGACCATCACCTTGGAGGTGGAGA 3187

Qy      1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323
Db      3188 GCTCTGACACCATTGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAG 3247

Qy      1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383
Db      3248 ACCAGCAGCGTCTGATCTTTGCGGGCAAGCAGCTGGAGGATGGCCGCACTCTCGCGGACT 3307

Qy      1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443
Db      3308 ACAACATCCAGAAGGAGAGCACCCTTCACCTTGTTCTCCGCCTCAGAGGTGGTATGCAGA 3367

Qy      1444 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 1503
Db      3368 TCTTTGTAAAGACCCTGACTGGAAAAACCATAACCCTGGAGGTTGAGAGCTCGGACACCA 3427

Qy      1504 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1563
Db      3428 TCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCCGGACCAGCAGCGTC 3487

Qy      1564 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1623
Db      3488 TGATCTTCGCCGGCAAACAGCTGGAGGATGGCCGCACCCTAGCAGACTACAACATCCAAA 3547

Qy      1624 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 1677
Db      3548 AGGAGAGCACCCTCCACCTTGTGCTCCGTCTCCGTGGTGGTCAGTAAGTCATGG 3601
```

RESULT 2
US-08-746-822-1
; Sequence 1, Application US/08746822
; Patent No. 6020190
; GENERAL INFORMATION:
; APPLICANT: Peter H. Quail
; APPLICANT: Alan H. Christensen
; APPLICANT: Howard P. Hershey
; APPLICANT: Robert A. Sharrock
; APPLICANT: Thomas D. Sullivan
; TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeff Lloyd
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,822
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,092
; FILING DATE: 5-JUNE-1995
; ATTORNEY/AGENT INFORMATION:

```
;      NAME:  Lloyd, Jeff
;      REGISTRATION NUMBER:  35,589
;      REFERENCE/DOCKET NUMBER:  08/462,092
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3840 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1993..3591
US-08-746-822-1

  Query Match             51.8%;  Score 1135.6;  DB 3;  Length 3840;
  Best Local Similarity   81.5%;  Pred. No. 0;
  Matches 1315;  Conservative    0;  Mismatches  299;  Indels    0;  Gaps    0;

Qy         64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
Db       1988 TGCAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATCACCCTCGAGGTGGAGT 2047

Qy        124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
Db       2048 CGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCAG 2107

Qy        184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
Db       2108 ACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGGCGCACGCTTGCTGACT 2167

Qy        244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
Db       2168 ACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTCAGGGGAGGCATGCAGA 2227

Qy        304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
Db       2228 TCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCCTCTGACACCA 2287

Qy        364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
Db       2288 TTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCAGCGGC 2347

Qy        424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
Db       2348 TCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGA 2407

Qy        484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
Db       2408 AGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGA 2467

Qy        544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
Db       2468 CCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAACGTCA 2527

Qy        604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
Db       2528 AGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGGCTCATCTTTGCTG 2587

Qy        664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
Db       2588 GAAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGAAGGAGAGCACCC 2647

Qy        724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 783
Db       2648 TCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCA 2707

Qy        784 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
Db       2708 AGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAATGTCAAGGCCAAGATCC 2767

Qy        844 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 903
Db       2768 AGGACAAGGAGGGCATCCCACCGGACCAGCAGCGTTTGATCTTCGCTGGCAAGCAGCTGG 2827

Qy        904 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 963
Db       2828 AGGATGGCCGCACCCTTGCGGATTACAACATCCAGAAGGAGAGCACCCTCCACCTGGTGC 2887

Qy        964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
Db       2888 TCCGTCTCAGGGGTGGTATGCAGATCTTTGTGAAGACACTCACTGGCAAGACAATCACCC 2947

Qy       1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083
Db       2948 TTGAGGTGGAGTCTTCGGATACCATTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGG 3007

Qy       1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143
Db       3008 GCATCCCACCCGACCAGCAGCGCCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCA 3067

Qy       1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203
Db       3068 CCCTGGCGGATTACAACATCCAGAAGGAGAGCACTCTCCACCTGGTGCTCCGCCTCAGGG 3127

Qy       1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263
Db       3128 GTGGCATGCAGATTTTTGTGAAGACATTGACTGGCAAGACCATCACCTTGGAGGTGGAGA 3187

Qy       1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323
Db       3188 GCTCTGACACCATTGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAG 3247

Qy       1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383
Db       3248 ACCAGCAGCGTCTGATCTTTGCGGGCAAGCAGCTGGAGGATGGCCGCACTCTCGCGGACT 3307

Qy       1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443
Db       3308 ACAACATCCAGAAGGAGAGCACCCTTCACCTTGTTCTCCGCCTCAGAGGTGGTATGCAGA 3367

Qy       1444 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 1503
Db       3368 TCTTTGTAAAGACCCTGACTGGAAAAACCATAACCCTGGAGGTTGAGAGCTCGGACACCA 3427

Qy       1504 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1563
Db       3428 TCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTC 3487

Qy       1564 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1623
Db       3488 TGATCTTCGCCGGCAAACAGCTGGAGGATGGCCGCACCCTAGCAGACTACAACATCCAAA 3547

Qy       1624 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 1677
Db       3548 AGGAGAGCACCCTCCACCTTGTGCTCCGTCTCCGTGGTGGTCAGTAAGTCATGG 3601


RESULT 3
US-09-094-350-1
; Sequence 1, Application US/09094350
; Patent No. 6054574
;  GENERAL INFORMATION:
;    APPLICANT:  Peter H. Quail
;    APPLICANT:  Alan H. Christensen
;    APPLICANT:  Howard P. Hershey
;    APPLICANT:  Robert A. Sharrock
;    APPLICANT:  Thomas D. Sullivan
;    TITLE OF INVENTION:  PLANT UBIQUITIN PROMOTER SYSTEM
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jeff Lloyd
```

```
;       STREET:  2421 N.W. 41st Street, Suite A-1
;       CITY:  Gainesville
;       STATE:  FL
;       COUNTRY:  USA
;       ZIP:  32606
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/094,350
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/746,822
;       FILING DATE:
;       APPLICATION NUMBER:  US/08/462,092
;       FILING DATE:  5-JUNE-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lloyd, Jeff
;       REGISTRATION NUMBER:  35,589
;       REFERENCE/DOCKET NUMBER:  08/462,092
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  904-375-8100
;       TELEFAX:  904-372-5800
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3840 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1993..3591
US-09-094-350-1

  Query Match             51.8%;  Score 1135.6;  DB 3;  Length 3840;
  Best Local Similarity   81.5%;  Pred. No. 0;
  Matches 1315;  Conservative    0;  Mismatches  299;  Indels    0;  Gaps    0;

Qy         64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
              ||   |||||||||||| ||||| || |||||||| ||||| ||||||||||||||| |||
Db       1988 TGCAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATCACCCTCGAGGTGGAGT 2047

Qy        124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
              |   ||||||||| || || || ||||| |||||||| || ||||| |||||||| || |
Db       2048 CGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCAG 2107

Qy        184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
              |||||||| |||| |||||||||||| |||||||| || || |||||||| ||   ||||||
Db       2108 ACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGGCGCACGCTTGCTGACT 2167

Qy        244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
              |||||||||||| |||  |||||| ||||| |||||||||||||| || || ||||| |
Db       2168 ACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTCAGGGGAGGCATGCAGA 2227

Qy        304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
              |||| ||||| || || || |||||||| |||||||| ||||||||| ||  ||||||||
Db       2228 TCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCCTCTGACACCA 2287

Qy        364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
              | || |||||||| || ||||||||||||||||| ||||| ||||| |||||||||  ||
Db       2288 TTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCAGCGGC 2347

Qy        424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
              | |||||||| || |||||||| || || |||||||| ||  | ||||||||||||||||
Db       2348 TCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGA 2407

Qy        484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
              | |||   ||||| ||| ||||||| || ||||| || || ||||||||||||||||||
```

►

```
Db       2408 AGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGA 2467

Qy        544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
              ||||||| || ||||| |||||||||||||||| |    ||||||||||| || ||||
Db       2468 CCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAACGTCA 2527

Qy        604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
              |||| |||||||| || ||||| |||||||| || || |||||| || | ||||||||  |
Db       2528 AGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGGCTCATCTTTGCTG 2587

Qy        664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
              |||| ||||| || || || || || ||  | ||||||||||||||||| |||  ||||
Db       2588 GAAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGAAGGAGAGCACCC 2647

Qy        724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 783
              | ||| |||| || || ||||| || || ||||| |||||||||||||| || || ||||
Db       2648 TCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCA 2707

Qy        784 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
              |||| |||||||| |||| |||  |||| |||| || ||||| || |||||||
Db       2708 AGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAATGTCAAGGCCAAGATCC 2767

Qy        844 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 903
              | |||||||| |||| || || ||||||||| | |||||||| || || ||||||||||
Db       2768 AGGACAAGGAGGGCATCCCACCGGACCAGCAGCGTTTGATCTTCGCTGGCAAGCAGCTGG 2827

Qy        904 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 963
              | ||||| |||||||  | || |||||||||||||| |||   ||||| ||||||||||
Db       2828 AGGATGGCCGCACCCTTGCGGATTACAACATCCAGAAGGAGAGCACCCTCCACCTGGTGC 2887

Qy        964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
              |||||||||| ||||| |||||||||||| |||||||| || ||||| |||||||||||
Db       2888 TCCGTCTCAGGGGTGGTATGCAGATCTTTGTGAAGACACTCACTGGCAAGACAATCACCC 2947

Qy       1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083
              | || |||||| |    || |||||||| |||||||||| |||||||| |||||||| |
Db       2948 TTGAGGTGGAGTCTTCGGATACCATTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGG 3007

Qy       1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143
              ||||||| || ||||||||| |  | ||||| ||||| || |||||||| ||||| || |
Db       3008 GCATCCCACCCGACCAGCAGCGCCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCA 3067

Qy       1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203
              ||||| | || |||||||||||||| ||| ||| | |||||||||||||| ||||| |
Db       3068 CCCTGGCGGATTACAACATCCAGAAGGAGAGCACTCTCCACCTGGTGCTCCGCCTCAGGG 3127

Qy       1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263
              |||| |||||||| || |||||||| ||||||| |||||||||| | |||||||||
Db       3128 GTGGCATGCAGATTTTTGTGAAGACATTGACTGGCAAGACCATCACCTTGGAGGTGGAGA 3187

Qy       1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323
                 |||||||||||| ||||| ||||| |||||||| |||||||| ||||| || || |
Db       3188 GCTCTGACACCATTGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAG 3247

Qy       1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383
              | |||||| | |||||||||| || || |||||||| ||||| ||||| ||  | ||||
Db       3248 ACCAGCAGCGTCTGATCTTTGCGGGCAAGCAGCTGGAGGATGGCCGCACTCTCGCGGACT 3307

Qy       1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443
              |||||||||||| |||  |||||| ||||| || ||||| || |||||||| |||||||
Db       3308 ACAACATCCAGAAGGAGAGCACCCTTCACCTTGTTCTCCGCCTCAGAGGTGGTATGCAGA 3367

Qy       1444 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 1503
              |||| || |||||||||||||| || ||||| || || || || |||      |||||||
Db       3368 TCTTTGTAAAGACCCTGACTGGAAAAACCATAACCCTGGAGGTTGAGAGCTCGGACACCA 3427

Qy       1504 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1563
              | || ||||| ||||| |||||||| |||||||| |||||||| || |||||||||  |
Db       3428 TCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTC 3487

Qy       1564 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1623
              ||||||| || || |||||||||| ||||| ||||||||   | ||||||||||||| |
Db       3488 TGATCTTCGCCGGCAAACAGCTGGAGGATGGCCGCACCCTAGCAGACTACAACATCCAAA 3547
```

```
Qy       1624 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 1677
Db       3548 AGGAGAGCACCCTCCACCTTGTGCTCCGTCTCCGTGGTGGTCAGTAAGTCATGG 3601


RESULT 4
5510474-1
;Patent No. 5510474
;    APPLICANT: QUAIL, PETER H.;CHRISTENSEN, ALAN H.;HERSHEY,
;HOWARD P.;SHARROCK, ROBERT A.;SULLIVAN, THOMAS D.
;    TITLE OF INVENTION: PLANT UBIQUITIN PROMOTER SYSTEM
;    NUMBER OF SEQUENCES: 12
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/296,268
;      FILING DATE: 25-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 191,134
;      FILING DATE: 03-FEB-1994
;      APPLICATION NUMBER: 76,363
;      FILING DATE: 11-JUN-1993
;      APPLICATION NUMBER: 670,496
;      FILING DATE: 15-MAR-1991
;      APPLICATION NUMBER: 194,824
;      FILING DATE: 17-MAY-1988
;SEQ ID NO:1:
;      LENGTH: 3840
5510474-1

  Query Match             51.7%;  Score 1134;  DB 9;  Length 3840;
  Best Local Similarity   81.4%;  Pred. No. 0;
  Matches 1314;  Conservative    0;  Mismatches  300;  Indels    0;  Gaps    0;

Qy         64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
Db       1988 TGCAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATCACCCTCGAGGTGGAGT 2047

Qy        124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
Db       2048 CGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCAG 2107

Qy        184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
Db       2108 ACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGGCGCACGCTTGCTGACT 2167

Qy        244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
Db       2168 ACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTCAGGGGAGGCATGCAGA 2227

Qy        304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
Db       2228 TCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCCTCTGACACCA 2287

Qy        364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
Db       2288 TTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCAGCGGC 2347

Qy        424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
Db       2348 TCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGA 2407

Qy        484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
Db       2408 AGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGA 2467

Qy        544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
Db       2468 CCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAACGTCA 2527

Qy        604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
Db       2528 AGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGGCTCATCTTTGCTG 2587

Qy        664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
Db       2588 GAAAGCAGCTTGAGGACGGGCGCACGCTTGCCGACTACAACATCCAGAAGGAGAGCACCC 2647

Qy        724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 783
Db       2648 TCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCA 2707

Qy        784 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
Db       2708 AGACTATCACCCTCGAGGTGGAGTCTTCAGACACCATCGACAATGTCAAGGCCAAGATCC 2767

Qy        844 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 903
Db       2768 AGGACAAGGAGGGCATCCCACCGGACCAGCAGCGTTTGATCTTCGCTGGCAAGCAGCTGG 2827

Qy        904 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 963
Db       2828 AGGATGGCCGCACCCTTGCGGATTACAACATCCAGAAGGAGAGCACCCTCCAGCTGGTGC 2887

Qy        964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
Db       2888 TCCGTCTCAGGGGTGGTATGCAGATCTTTGTGAAGACACTCACTGGCAAGACAATCACCC 2947

Qy       1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083
Db       2948 TTGAGGTGGAGTCTTCGGATACCATTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGG 3007

Qy       1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143
Db       3008 GCATCCCACCCGACCAGCAGCGCCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCA 3067

Qy       1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203
Db       3068 CCCTGGCGGATTACAACATCCAGAAGGAGAGCACTCTCCACCTGGTGCTCCGCCTCAGGG 3127

Qy       1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263
Db       3128 GTGGCATGCAGATTTTTGTGAAGACATTGACTGGCAAGACCATCACCTTGGAGGTGGAGA 3187

Qy       1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323
Db       3188 GCTCTGACACCATTGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAG 3247

Qy       1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383
Db       3248 ACCAGCAGCGTCTGATCTTTGCGGGCAAGCAGCTGGAGGATGGCCGCACTCTCGCGGACT 3307

Qy       1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443
Db       3308 ACAACATCCAGAAGGAGAGCACCCTTCACCTTGTTCTCCGCCTCAGAGGTGGTATGCAGA 3367

Qy       1444 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 1503
Db       3368 TCTTTGTAAAGACCCTGACTGGAAAAACCATAACCCTGGAGGTTGAGAGCTCGGACACCA 3427

Qy       1504 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1563
Db       3428 TCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTC 3487

Qy       1564 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1623
Db       3488 TGATCTTCGCCGGCAAACAGCTGGAGGATGGCCGCACCCTAGCAGACTACAACATCCAAA 3547

Qy       1624 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 1677
Db       3548 AGGAGAGCACCCTCCACCTTGTGCTCCGTCTCCGTGGTGGTCAGTAAGTCATGG 3601


RESULT 5
US-09-516-277-1
; Sequence 1, Application US/09516277
; Patent No. 6528701
```

```
; GENERAL INFORMATION:
;  APPLICANT: WANG -- OARD
;  TITLE OF INVENTION: RICE UBIQUITIN-DERIVED  PROMOTERS
;  FILE REFERENCE: 98A5
;  CURRENT APPLICATION NUMBER: US/09/516,277
;  CURRENT FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3184
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1702)..(3072)
US-09-516-277-1

  Query Match             45.2%;  Score 991.4;  DB 3;  Length 3184;
  Best Local Similarity   82.2%;  Pred. No. 4.4e-278;
  Matches 1139;  Conservative    0;  Mismatches  246;  Indels    0;  Gaps    0;

Qy          64 TGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGC 123
               ||  ||||||||||| ||||||| || || || ||||||||||||||||||||||||
Db        1697 TGCAGATGCAGATCTTTGTGAAGACCCTCACCGGCAAGACCATCACCCTCGAGGTTGAGT 1756

Qy         124 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 183
               ||   |||||||| || ||||||||||| |||||||| || ||||| |||||||| || |
Db        1757 CCTCGGACACCATTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCGG 1816

Qy         184 ACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACT 243
               |||||||| | || ||||| ||||| || ||||| || ||||| ||||||||| | ||||
Db        1817 ACCAGCAGCGTCTCATCTTCGCTGGCAAGCAGCTTGAGGATGGCCGCACCCTGGCCGACT 1876

Qy         244 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAA 303
               |||||||||||| |||||||||| ||||| |||||| | ||||| || || |||||| |
Db        1877 ACAACATCCAGAAGGAGTCCACCCTCCACCTTGTGCTCAGGCTCAGGGGAGGCATGCAGA 1936

Qy         304 TCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCA 363
               ||||||| |||||   | |||||||||||||||||||||||||| ||| |   ||||||||
Db        1937 TCTTCGTCAAGACCTTGACTGGCAAGACCATCACCCTTGAGGTCGAGTCGTCTGACACCA 1996

Qy         364 TCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGT 423
               | || || ||||| || |||||||||||||||| ||||| || || ||||||||| |
Db        1997 TTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTC 2056

Qy         424 TGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 483
               | ||||| || || |||||||| || ||||| |||||||||| ||||||||||||||||||
Db        2057 TCATCTTCGCTGGCAAGCAGCTTGAGGATGGCCGCACCCTGGCTGACTACAACATCCAGA 2116

Qy         484 AAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 543
               | ||||| ||||| ||||| |||||| | ||||| || || ||||||||||||||| ||||
Db        2117 AGGAGTCCACCCTCCACCTTGTGCTCAGGCTCAGGGGAGGCATGCAGATCTTCGTCAAGA 2176

Qy         544 CCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCA 603
               || ||||||| |||||||||||||||||| ||| |   |||||||| || |||||||
Db        2177 CCTTGACTGGCAAGACCATCACCCTCGAGGTCGAGTCGTCTGACACCATTGACAATGTCA 2236

Qy         604 AGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCG 663
               |||| |||||||| || ||||| ||||| || || || |||||| |   | ||||| ||||
Db        2237 AGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCAGACCAGCAGCGTCTCATCTTCGCCG 2296

Qy         664 GAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT 723
               | || |||||||| ||||| || |||||  ||||||||||||||||||| |||||||||
Db        2297 GCAAGCAGCTGGAGGATGGCCGCACCCTTGCTGACTACAACATCCAGAAGGAGTCCACCC 2356

Qy         724 TGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 783
               | ||||| || ||| | ||||| || || ||||| |||||||| ||||| || || ||||
Db        2357 TCCACCTTGTGCTCAGGCTCAGGGGAGGTATGCAGATCTTCGTCAAGACCCTGACCGGCA 2416

Qy         784 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
               |||||||||||| ||||||||| ||   ||||| ||||| || || ||||| ||||||||
Db        2417 AGACCATCACCCTCGAGGTCGAGTCCTCGGACACGATCGACAATGTGAAAGCCAAGATCC 2476

Qy         844 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 903
               | |||||||| ||||| || || ||||||||| |   | ||||||| || ||||||||||
Db        2477 AGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATCTTTGCTGGCAAGCAGCTGG 2536

Qy         904 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 963
               | ||||| |||||||   | |||||||||||||||| ||||| ||||| ||||| ||||
Db        2537 AGGATGGCCGCACCCTTGCCGACTACAACATCCAGAAGGAGTCCACCCTCCACCTTGTGC 2596

Qy         964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
               || | ||||| ||||| |||||||||||||| |||||||||| || |||||||||| |
Db        2597 TCAGGCTCAGGGGTGGTATGCAGATCTTCGTCAAGACCCTGACCGGCAAGACCATCACGC 2656

Qy        1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083
               | || || ||| |     ||||| || || ||||| ||||| |||||||| |||||||| |
Db        2657 TTGAGGTCGAGTCCTCGGACACGATCGACAATGTGAAGGCCAAGATCCAGGACAAGGAGG 2716

Qy        1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143
               | ||||| || ||||||||| |   | ||||| ||||| || ||||| || ||||| || |
Db        2717 GTATCCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTTGAGGATGGCCGCA 2776

Qy        1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203
               || || |||||||||||||||||| ||||||||| | |||||||| ||| | |||| |
Db        2777 CCTTGGCTGACTACAACATCCAGAAGGAGTCCACCCTTCACCTGGTTCTCAGGCTCAGGG 2836

Qy        1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263
               |||||||||||||||||||||||||||||||||| |||||||| || || ||||| |||
Db        2837 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGCAAGACCATTACCCTTGAGGTTGAGT 2896

Qy        1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323
               ||   ||||| ||||| || || ||||| |||||||| |||||||| |||||||| || |
Db        2897 CGTCCGACACTATTGACAACGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCGG 2956

Qy        1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383
               | |||||| |  |||||||||||| |||||| || ||||| |||||| || | || |
Db        2957 ACCAGCAGCGTCTGATCTTTGCTGGTAAGCAGCTTGAGGATGGCCGCACCTTGGCGGATT 3016

Qy        1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443
               |||||||||||| |||||||| || |||||||| || || ||  | |||||   | |
Db        3017 ACAACATCCAGAAGGAGTCCACACTCCACCTGGTTCTGCGCCTCCGTGGTGGCCAGTAAG 3076

Qy        1444 TCTTC 1448
               || ||
Db        3077 TCCTC 3081


RESULT 6
US-09-516-277-3
; Sequence 3, Application US/09516277
; Patent No. 6528701
; GENERAL INFORMATION:
;  APPLICANT: WANG -- OARD
;  TITLE OF INVENTION: RICE UBIQUITIN-DERIVED  PROMOTERS
;  FILE REFERENCE: 98A5
;  CURRENT APPLICATION NUMBER: US/09/516,277
;  CURRENT FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 4442
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2786)..(4159)
US-09-516-277-3

  Query Match             45.1%;  Score 988.8;  DB 3;  Length 4442;
  Best Local Similarity   82.2%;  Pred. No. 3e-277;
  Matches 1137;  Conservative    0;  Mismatches  247;  Indels    0;  Gaps    0;

Qy          52 GTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCC 111
               |||||  |  ||||   |||||||||| ||||||||  |||| || ||||| |||||||
Db        2769 GTGATTATTTCTTGCAGATGCAGATCTTTGTGAAGACATTGACCGGCAAGACTATCACCC 2828
```

```
Qy        112 TCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 171
              ||||||| ||| ||  |||||||||||| |||||||||| |||||||||||||||||| |
Db       2829 TCGAGGTGGAGTCCTCTGACACCATCGATAATGTCAAGGCTAAGATCCAAGATAAGGAGG 2888

Qy        172 GCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCA 231
              ||||||| || |||||||| | || ||||| ||||| || |||||||| |||||  | |
Db       2889 GCATCCCCCCGGACCAGCAGCGTCTCATCTTCGCTGGCAAGCAGCTGGAGGATGGCAGGA 2948

Qy        232 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 291
              ||||  |||||||||||||||||| ||||| ||||| ||||| || ||||| ||| | |
Db       2949 CCCTTGCTGACTACAACATCCAGAAGGAGTCGACCCTTCACCTTGTCCTCCGCCTCCGTG 3008

Qy        292 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGC 351
              |||| ||||| ||||| || ||||| || || |||||||| ||||||||||||||||||
Db       3009 GTGGCATGCAGATCTTTGTCAAGACTCTGACCGGCAAGACTATCACCCTTGAGGTGGAGT 3068

Qy        352 CCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTG 411
              |   |||||||||||| |||||||| || |||||||||||||| || |||||| || || |
Db       3069 CTTCTGACACCATCGACAACGTCAAGGCCAAGATCCAGGACAAAGAGGGCATCCCCCCAG 3128

Qy        412 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 471
              |||||||| |  | ||||| ||||| |||||||||| ||||| | |||||  ||||||
Db       3129 ACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCAGGACCCTTGCTGACT 3188

Qy        472 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 531
              |||||||||||| ||||| ||||| ||||| || ||||| ||| | ||||| |||||||
Db       3189 ACAACATCCAGAAGGAGTCCACCCTCCACCTTGTCCTCCGCCTCCGTGGTGGCATGCAGA 3248

Qy        532 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCA 591
              |||| || ||||| ||||| || |||||||||||||||||||||| |   ||||||||
Db       3249 TCTTTGTCAAGACACTGACCGGCAAGACCATCACCCTCGAGGTGGAATCTTCTGACACCA 3308

Qy        592 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGT 651
              |||| || |||||||| |||||||| || ||||| |||||||| || || |||||| |
Db       3309 TCGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCGGACCAGCAGCGTC 3368

Qy        652 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 711
              | |||||||||| || ||||| || || ||  | |||||  ||||||||||||||||||
Db       3369 TCATCTTTGCCGGCAAGCAGCTTGAGGACGGCAGGACCCTTGCTGACTACAACATCCAGA 3428

Qy        712 AAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 771
              | ||||| ||  | ||||| || |||||||||| || || ||||||||||||||||||
Db       3429 AGGAGTCAACGCTTCACCTTGTCCTCCGTCTCAGGGGAGGCATGCAAATCTTCGTGAAGA 3488

Qy        772 CACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCA 831
              | || || |||||||||||||||||| ||||| ||| |   ||| || |||||  || ||||
Db       3489 CTCTGACCGGCAAGACCATCACCCTCGAGGTGGAGTCTTCTGATACCATCGACAATGTCA 3548

Qy        832 AAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 891
              | || |||||||| |||||||| |||||||| || |||||||||  |  ||||||||  |
Db       3549 AGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCGGACCAGCAGCGCCTCATCTTTGCTG 3608

Qy        892 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 951
              | |||||||||| |||||  | |||||  ||||||||||||||||||| ||||| ||||
Db       3609 GCAAGCAGCTGGAGGATGGCAGGACCCTTGCTGACTACAACATCCAGAAGGAGTCCACCC 3668

Qy        952 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1011
              | ||||| |||||||| ||  | ||||| |||||||||| || |||||||| || || |
Db       3669 TCCACCTTGTGCTCCGCCTTCGTGGTGGTATGCAGATCTTTGTCAAGACCCTCACAGGCA 3728

Qy       1012 AGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1071
              |||||||||||| || || |||       |||||||| || || |||||||| |||||||
Db       3729 AGACCATCACCCTGGAGGTTGAGAGCTCGGACACCATCGACAACGTCAAGGCCAAGATCC 3788

Qy       1072 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 1131
              | |||||||| |||||||| || ||||||||  |  | ||||| ||||| || ||||| |
Db       3789 AGGACAAGGAGGGCATCCCCCCAGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTCG 3848

Qy       1132 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGC 1191
              | ||||| || |||||  | |||||||||||||||||| |||||| ||| | ||||||||||
Db       3849 AGGATGGCCGCACCCTTGCCGACTACAACATCCAGAAGGAGTCTACCCTCCACCTGGTGC 3908

Qy       1192 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1251
              | |||||| | ||||| |||||||||||||||||||| ||||||| ||||||||||||
Db       3909 TTCGTCTCCGTGGTGGTATGCAGATCTTCGTGAAGACCTTGACTGGGAAGACCATCACTT 3968

Qy       1252 TCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1311
              | ||||| |||       |||||||||| ||||| ||||| |||||||| |||||||| |
Db       3969 TGGAGGTTGAGAGCTCCGACACCATTGATAATGTGAAGGCCAAGATCCAGGACAAGGAGG 4028

Qy       1312 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1371
              | || || || || |||||| |  ||||||| ||||| || |||||||| ||||||||||
Db       4029 GGATTCCCCCAGACCAGCAGCGTCTGATCTTCGCTGGCAAGCAGCTGGAGGATGGACGCA 4088

Qy       1372 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAG 1431
              ||||  | ||||||||||||||||| |||||||||| |||||||||||||| ||  | |
Db       4089 CCCTCGCCGACTACAACATCCAGAAGGAGTCCACCCTCCACCTGGTGCTCCGCCTCCGTG 4148

Qy       1432 GTGG 1435
              ||||
Db       4149 GTGG 4152
```

RESULT 7
US-09-866-153-8
; Sequence 8, Application US/09866153
; Patent No. 6638766
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/866,153
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5125)
; OTHER INFORMATION: The "n" at position 5125 is any nucleotide.
; NAME/KEY: misc_feature
; LOCATION: (5150)
; OTHER INFORMATION: The "n" at position 5150 is any nucleotide.
US-09-866-153-8

Query Match 38.0%; Score 832.4; DB 3; Length 5174;
Best Local Similarity 82.1%; Pred. No. 1.1e-231;
Matches 947; Conservative 7; Mismatches 200; Indels 0; Gaps 0;

```
Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583
              ||||||||| ||||| |||||||| ||||| |||||||||||||| ||||| ||| ||
Db       3688 GATGCAGATTTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCCTC 3747

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643
```

```
Db       3748 AGACACTATTGACAATGTCAAGGCTAAGATCCAGGACAAGGAAGGCATTCCTCCAGATCA 3807

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703
              ||||||| ||||:||||| || || ||||| || ||||| ||||||||  ||||||||||
Db       3808 GCAGAGGCTGATYTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGCTGACTACAA 3867

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763
              |||||||| |||||||| | |||||||| ||| | || || || || ||||| || ||
Db       3868 CATCCAGAAGGAGTCCACCCTCCACCTGGTGCTCAGGCTTAGGGGAGGCATGCAGATTTT 3927

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
              ||| ||||| |||||||||||||| ||||| |||||||||||  |    |||||| |||||
Db       3928 CGTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTCGAGTCTTCTGACACGATCGA 3987

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883
              ||||| || || |||||||| |||||||| || || || || ||||||||| | :| ||
Db       3988 CAACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCGGACCAGCAGCGTYTCAT 4047

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943
              || || || |||||||| || ||||| |||||||| ||||||||||||||||||| ||
Db       4048 TTTCGCTGGCAAGCAGCTCGAGGATGGCCGCACCCTCGCTGACTACAACATCCAGAAGGA 4107

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003
              ||| || || ||||| |||||| | ||||| ||||| |||||||||||||| ||||||||
Db       4108 GTCGACTCTCCACCTTGTGCTCAGGCTCAGGGGTGGCATGCAGATCTTCGTCAAGACCCT 4167

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063
              ||||| |||||||||| | || |||||| |    |||||||||| ||||| |||||
Db       4168 CACTGGCAAGACCATCACCTTGGAGGTGGAGTCCTCGGACACCATTGACAATGTGAAGGC 4227

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              |||||||| |||||||| |||||||| || ||||||||| |   | ||:|| ||||| ||
Db       4228 GAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATYTTCGCCGGCAA 4287

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183
              ||:|| || ||||| || |||||  | || ||||||||||||||:||||||||| | ||
Db       4288 GCARCTTGAGGATGGCCGCACCCTTGCGGANTACAACATCCAGAARGAGTCCACCCTTCA 4347

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243
              |||||||||||| ||  | ||||| |||||||| ||||| |||||||| || || |||||
Db       4348 CCTGGTGCTCCGCCTTCGTGGTGGTATGCAGATTTTCGTCAAGACCCTCACCGGCAAGAC 4407

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303
              |||||| || ||||||||| |   |||||||||||| ||||| ||||| |||||||| ||
Db       4408 CATCACCCTGGAGGTGGAGTCCTCTGACACCATTGACAATGTGAAGGCGAAGATCCAGGA 4467

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363
              ||||| |||||||| || || |||||| | :| |||||||||| || ||||| || ||
Db       4468 TAAGGAGGGCATCCCCCCGGACCAGCAGCGTYTTATCTTTGCTGGCAAGCAGCTTGAGGA 4527

Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423
              ||| |||||||||| | || |||||||||||||| |||||||||| ||||||||||||||
Db       4528 TGGCCGCACCCTGGCAGANTACAACATCCAGAAGGAGTCCACCCTTCACCTGGTGCTCCG 4587

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1483
              ||| | ||||| ||||||||:||||| |||||||| || || |||||||||| || ||
Db       4588 CCTTCGCGGTGGTATGCAGATYTTCGTCAAGACCCTCACCGGCAAGACCATCACCCTGGA 4647

Qy       1484 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1543
              |||||| |   ||||||||| || ||||| ||||| |||||||| |||||||| |||||
Db       4648 GGTGGAGTCCTCTGACACCATCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCAT 4707

Qy       1544 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1603
              ||| || |||||||| |  | ||||| || || || |||||||| ||||| ||||||||
Db       4708 CCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCACCCT 4767

Qy       1604 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1663
              | | |||||||||||||||| |||||||| || |||||||||||||||||  | |||||
Db       4768 GGCAGACTACAACATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCGTCTCCGTGGTGG 4827

Qy       1664 GATGCAGATCTTCG 1677
                 | |  || | |
Db       4828 CCAGTAAGTCCTGG 4841
```

RESULT 8
US-09-693-467A-8
; Sequence 8, Application US/09693467A
; Patent No. 6686513
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-04331
; CURRENT APPLICATION NUMBER: US/09/693,467A
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/270,976
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5125)
; OTHER INFORMATION: The "n" at position 5125 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5150)
; OTHER INFORMATION: The "n" at position 5150 is any nucleotide.
US-09-693-467A-8

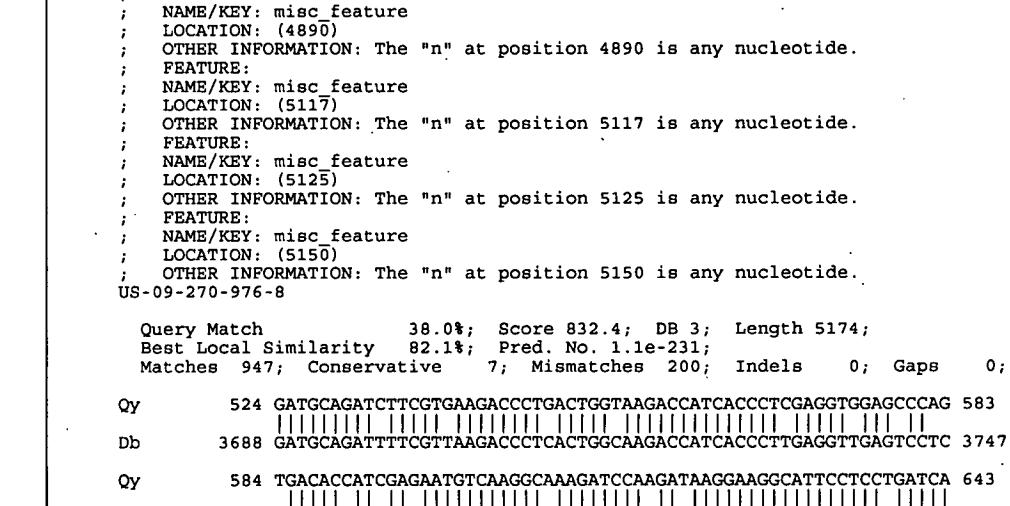
Query Match 38.0%; Score 832.4; DB 3; Length 5174;
Best Local Similarity 82.1%; Pred. No. 1.1e-231;
Matches 947; Conservative 7; Mismatches 200; Indels 0; Gaps 0;

```
Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583
              ||||||||| ||||| |||||||| ||||| |||||||||||||| ||||| ||| ||
Db       3688 GATGCAGATTTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCCTC 3747

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643
              ||||| || || |||||||||| |||||||| || ||||||||||||||||||| |||||
Db       3748 AGACACTATTGACAATGTCAAGGCTAAGATCCAGGACAAGGAAGGCATTCCTCCAGATCA 3807

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703
              ||||||| ||||:||||| || || ||||| || ||||| ||||||||  ||||||||||
Db       3808 GCAGAGGCTGATYTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGCTGACTACAA 3867

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763
              |||||||| |||||||| | |||||||| ||| | || || || || ||||| || ||
Db       3868 CATCCAGAAGGAGTCCACCCTCCACCTGGTGCTCAGGCTTAGGGGAGGCATGCAGATTTT 3927

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
```

```
Db       3928 CGTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTCGAGTCTTCTGACACGATCGA 3987

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883

Db       3988 CAACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCGGACCAGCAGCGTYTCAT 4047

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943

Db       4048 TTTCGCTGGCAAGCAGCTCGAGGATGGCCGCACCCTCGCTGACTACAACATCCAGAAGGA 4107

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003

Db       4108 GTCGACTCTCCACCTTGTGCTCAGGCTCAGGGGTGGCATGCAGATCTTCGTCAAGACCCT 4167

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063

Db       4168 CACTGGCAAGACCATCACCTTGGAGGTGGAGTCCTCGGACACCATTGACAATGTGAAGGC 4227

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123

Db       4228 GAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATYTTCGCCGGCAA 4287

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183

Db       4288 GCARCTTGAGGATGGCCGCACCCTTGCGGANTACAACATCCAGAARGAGTCCACCCTTCA 4347

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243

Db       4348 CCTGGTGCTCCGCCTTCGTGGTGGTATGCAGATTTTCGTCAAGACCCTCACCGGCAAGAC 4407

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303

Db       4408 CATCACCCTGGAGGTGGAGTCCTCTGACACCATTGACAATGTGAAGGCGAAGATCCAGGA 4467

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363

Db       4468 TAAGGAGGGCATCCCCCCGGACCAGCAGCGTYTTATCTTTGCTGGCAAGCAGCTTGAGGA 4527

Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423

Db       4528 TGGCCGCACCCTGGCAGANTACAACATCCAGAAGGAGTCCACCCTTCACCTGGTGCTCCG 4587

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1483

Db       4588 CCTTCGCGGTGGTATGCAGATYTTCGTCAAGACCCTCACCGGCAAGACCATCACCCTGGA 4647

Qy       1484 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1543

Db       4648 GGTGGAGTCCTCTGACACCATCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCAT 4707

Qy       1544 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1603

Db       4708 CCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCACCCT 4767

Qy       1604 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1663

Db       4768 GGCAGACTACAACATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCGTCTCCGTGGTGG 4827

Qy       1664 GATGCAGATCTTCG 1677

Db       4828 CCAGTAAGTCCTGG 4841
```

RESULT 9
US-09-270-976-8
; Sequence 8, Application US/09270976A
; Patent No. 6706948
; GENERAL INFORMATION:
; APPLICANT: Albert, Henrik H.
; APPLICANT: Wei, Hairong
; TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
; FILE REFERENCE: UH-03648
; CURRENT APPLICATION NUMBER: US/09/270,976A
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5174
; TYPE: DNA
; ORGANISM: Saccharum Hybrid Cultivar H32-8560
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: The "n" at position 9 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4318)
; OTHER INFORMATION: The "n" at position 4318 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4546)
; OTHER INFORMATION: The "n" at position 4546 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4890)
; OTHER INFORMATION: The "n" at position 4890 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5117)
; OTHER INFORMATION: The "n" at position 5117 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5125)
; OTHER INFORMATION: The "n" at position 5125 is any nucleotide.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5150)
; OTHER INFORMATION: The "n" at position 5150 is any nucleotide.
US-09-270-976-8

Query Match 38.0%; Score 832.4; DB 3; Length 5174;
Best Local Similarity 82.1%; Pred. No. 1.1e-231;
Matches 947; Conservative 7; Mismatches 200; Indels 0; Gaps 0;



```
Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583

Db       3688 GATGCAGATTTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCCTC 3747

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643

Db       3748 AGACACTATTGACAATGTCAAGGCTAAGATCCAGGACAAGGAAGGCATTCCTCCAGATCA 3807

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703

Db       3808 GCAGAGGCTGATYTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGCTGACTACAA 3867

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763

Db       3868 CATCCAGAAGGAGTCCACCCTCCACCTGGTGCTCAGGCTTAGGGGAGGCATGCAGATTTT 3927

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823

Db       3928 CGTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTCGAGTCTTCTGACACGATCGA 3987

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883

Db       3988 CAACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCGGACCAGCAGCGTYTCAT 4047

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943

Db       4048 TTTCGCTGGCAAGCAGCTCGAGGATGGCCGCACCCTCGCTGACTACAACATCCAGAAGGA 4107

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003

Db       4108 GTCGACTCTCCACCTTGTGCTCAGGCTCAGGGGTGGCATGCAGATCTTCGTCAAGACCCT 4167
```

```
Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063
             ||||| |||||||||||| | || |||||| |    |||||||||| ||||| |||||
Db       4168 CACTGGCAAGACCATCACCTTGGAGGTGGAGTCCTCGGACACCATTGACAATGTGAAGGC 4227

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              |||||||| |||||||| |||||||| || ||||||||  |   | ||:|| ||||| ||
Db       4228 GAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATYTTCGCCGGCAA 4287

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183
              ||:|| || ||||| || |||||  | || ||||||||||||||:||||||||| | ||
Db       4288 GCARCTTGAGGATGGCCGCACCCTTGCGGANTACAACATCCAGAARGAGTCCACCCTTCA 4347

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243
              |||||||||||| ||  | ||||| |||||||| ||||| |||||||| || || |||||
Db       4348 CCTGGTGCTCCGCCTTCGTGGTGGTATGCAGATTTTCGTCAAGACCCTCACCGGCAAGAC 4407

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303
              |||||| || ||||||||| |   |||||||||||| ||||| ||||| |||||||| ||
Db       4408 CATCACCCTGGAGGTGGAGTCCTCTGACACCATTGACAATGTGAAGGCGAAGATCCAGGA 4467

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363
              ||||| |||||||| || || |||||| | :| |||||||||| || ||||| || ||
Db       4468 TAAGGAGGGCATCCCCCCGGACCAGCAGCGTYTTATCTTTGCTGGCAAGCAGCTTGAGGA 4527

Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423
              ||| ||||||||| | || |||||||||||||| |||||||||| ||||||||||||||
Db       4528 TGGCCGCACCCTGGCAGANTACAACATCCAGAAGGAGTCCACCCTTCACCTGGTGCTCCG 4587

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1483
              ||| | ||||| ||||||||:||||| |||||||| || || |||||||||| || ||
Db       4588 CCTTCGCGGTGGTATGCAGATYTTCGTCAAGACCCTCACCGGCAAGACCATCACCCTGGA 4647

Qy       1484 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1543
              |||||| |   |||||||| || ||||| ||||| |||||||| |||||||| ||||||
Db       4648 GGTGGAGTCCTCTGACACCATCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCAT 4707

Qy       1544 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1603
              ||| || |||||||||  |   | ||||| || || || |||||||| ||||| |||||||
Db       4708 CCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCACCCT 4767

Qy       1604 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1663
              | | ||||||||||||||||| |||||||| || |||||||||||||||   | |||||
Db       4768 GGCAGACTACAACATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCGTCTCCGTGGTGG 4827

Qy       1664 GATGCAGATCTTCG 1677
                | |  || | |
Db       4828 CCAGTAAGTCCTGG 4841


RESULT 10
US-09-866-153-5
; Sequence 5, Application US/09866153
; Patent No. 6638766
; GENERAL INFORMATION:
;  APPLICANT: Albert, Henrik H.
;  APPLICANT:  Wei, Hairong
;  TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
;  FILE REFERENCE: UH-03648
;  CURRENT APPLICATION NUMBER: US/09/866,153
;  CURRENT FILING DATE:  2001-05-24
;  PRIOR APPLICATION NUMBER: 09/270,976
;  PRIOR FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 5512
;   TYPE: DNA
;   ORGANISM: Saccharum Hybrid Cultivar H32-8560
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2218)
;   OTHER INFORMATION: The "n" at position 2218 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2227)
;   OTHER INFORMATION: The "n" at position 2227 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2229)
;   OTHER INFORMATION: The "n" at position 2229 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2249)
;   OTHER INFORMATION: The "n" at position 2249 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2282)
;   OTHER INFORMATION: The "n" at position 2282 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2327)
;   OTHER INFORMATION: The "n" at position 2327 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2341)
;   OTHER INFORMATION: The "n" at position 2341 is any nucleotide.
;   NAME/KEY: misc_feature
;   LOCATION: (2477)
;   OTHER INFORMATION: The "n" at position 2477 is any nucleotide.
US-09-866-153-5

  Query Match             29.8%;  Score 654.2;  DB 3;  Length 5512;
  Best Local Similarity   80.4%;  Pred. No. 7.9e-180;
  Matches  755;  Conservative    4;  Mismatches  180;  Indels    0;  Gaps    0;

Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583
              |||||||||||||||| |||||||| ||||| |||||||||||||| ||||| ||| |
Db       1802 GATGCAGATCTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCTTC 1861

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643
              ||||: || || :|||||:|||| ||||| || || |||||||||||||||| |||||
Db       1862 AGACAMTATTGACMATGTCMAGGCTAAGATACAGGACAAGGAAGGCATTCCTCCGGATCA 1921

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703
              ||||||| |||||||||| || || ||||| || ||||| |||||||| :|||||||||
Db       1922 GCAGAGGCTGATCTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGYTGACTACAA 1981

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763
              ||||||||| ||||||||| ||| | ||||| || || ||||||||||
Db       1982 CATCCAGAAGGAGTCCACCSTCCACCTGGTGCTCAGGCTCAGGGGAGGCATGCAAATCTT 2041

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
              ||| ||||| |||||||||||||| ||||| |||||||||||  |    |||||| |||||
Db       2042 CGTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTCGAGTCTTCTGACACGATCGA 2101

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883
              ||||| || || |||||||| |||||||| || || || || |||||||| |  | ||
Db       2102 CAACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCGGACCAGCAGCGTCTCAT 2161

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943
              ||| || || |||||||| || ||||| ||||||||   ||||||||||||||||| ||
Db       2162 CTTCGCTGGCAAGCAGCTCGAGGATGGCCGCACCCTCGCTGACTACAACATCCAGANGGA 2221

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003
              ||| |   | ||||| |||||| |  | || ||||| |||||||| || || |||||| |
Db       2222 GTCGANTNTCCACCTTGTGCTCAGGTTNAGGGGTGGCATGCAGATTTTTGTCAAGACCTT 2281

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063
              ||||| |||||||||||| | || |||||| |    |||||||| || ||||| ||||
Db       2282 NACTGGCAAGACCATCACCTTGGAGGTGGAGTCTTCGGACACCATNGACAATGTGAAGGN 2341

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              |||||||| |||||||||||| ||||| || |||||||||  |  | || ||||| || ||
Db       2342 GAAGATCCAGGACAAGGAAGGAATCCCCCCAGACCAGCAGCGTCTTATTTTTGCTGGCAA 2401

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183
              ||||| || ||||| || |||||  | |||||||||||||||| |||||||| | ||
Db       2402 GCAGCTTGAGGATGGCCGCACCCTAGCAGACTACAACATCCAGAAGGAGTCCACCCTTCA 2461

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243
              ||||||||||||  |  | ||||| |||||||||||||| |||||||| || || |||||
Db       2462 CCTGGTGCTCCGCCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 2521
```

```
Db       2462 CCTGGTGCTCCGCTTNCGCGGTGGTATGCAGATCTTCGTCAAGACCCTCACCGGCAAGAC 2521

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303
              |||||| || ||||||||| |    ||||||||| || ||||| ||||| ||||||||| ||
Db       2522 CATCACCCTGGAGGTGGAGTCCTCTGACACCATCGACAATGTGAAGGCGAAGATCCAGGA 2581

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363
              |||||| |||||||| || || |||||| |  | ||||| || || || |||||||| ||
Db       2582 CAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGA 2641

Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423
              ||| |||||||||  | ||||||||||||||||| |||||||| || ||||||||||||||
Db       2642 TGGCCGCACCCTGGCAGACTACAACATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCG 2701

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1462
              |||  | |||||   | |   ||  |  | ||  ||| |
Db       2702 TCTCCGTGGTGGCCAGTAATCCTGGGCCATGAAGCTGTC 2740


RESULT 11
US-09-693-467A-5
; Sequence 5, Application US/09693467A
; Patent No. 6686513
; GENERAL INFORMATION:
;  APPLICANT: Albert, Henrik H.
;  APPLICANT:  Wei, Hairong
;  TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
;  FILE REFERENCE: UH-04331
;  CURRENT APPLICATION NUMBER: US/09/693,467A
;  CURRENT FILING DATE:  2000-10-20
;  PRIOR APPLICATION NUMBER: 09/270,976
;  PRIOR FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 5512
;   TYPE: DNA
;   ORGANISM: Saccharum Hybrid Cultivar H32-8560
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2218)
;   OTHER INFORMATION: The "n" at position 2218 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2227)
;   OTHER INFORMATION: The "n" at position 2227 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2229)
;   OTHER INFORMATION: The "n" at position 2229 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2249)
;   OTHER INFORMATION: The "n" at position 2249 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2282)
;   OTHER INFORMATION: The "n" at position 2282 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2327)
;   OTHER INFORMATION: The "n" at position 2327 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2341)
;   OTHER INFORMATION: The "n" at position 2341 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2477)
;   OTHER INFORMATION: The "n" at position 2477 is any nucleotide.
US-09-693-467A-5

  Query Match             29.8%;  Score 654.2;  DB 3;  Length 5512;
  Best Local Similarity   80.4%;  Pred. No. 7.9e-180;
  Matches  755;  Conservative    4;  Mismatches  180;  Indels    0;  Gaps    0;

Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583
              |||||||||||||| |||||||| ||||| |||||||||||||| ||||| ||| |
Db       1802 GATGCAGATCTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCTTC 1861

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643
              ||||: || || :|||||:|||| ||||| || || ||||||||||||||||| |||||
Db       1862 AGACAMTATTGACMATGTCMAGGCTAAGATACAGGACAAGGAAGGCATTCCTCCGGATCA 1921

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703
              ||||||| |||||||||| || || ||||| || ||||| |||||||| :|||||||||
Db       1922 GCAGAGGCTGATCTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGYTGACTACAA 1981

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763
              ||||||||| ||||||||| | |||||||| ||| | ||||| || || ||||||||||
Db       1982 CATCCAGAAGGAGTCCACCSTCCACCTGGTGCTCAGGCTCAGGGGAGGCATGCAAATCTT 2041

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
              ||| ||||| |||||||||||||| ||||| |||||||||||| |   |||||| |||||
Db       2042 CGTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTCGAGTCTTCTGACACGATCGA 2101

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883
              ||||| || || |||||||| |||||||| || || || || ||||||||| |  | ||
Db       2102 CAACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCGGACCAGCAGCGTCTCAT 2161

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943
              ||| || || |||||||| || ||||| ||||||||  |||||||||||||||||||  ||
Db       2162 CTTCGCTGGCAAGCAGCTCGAGGATGGCCGCACCCTCGCTGACTACAACATCCAGANGGA 2221

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003
              ||| |   | ||||| |||||| |  | || ||||| |||||||| || || |||||| |
Db       2222 GTCGANTNTCCACCTTGTGCTCAGGTTNAGGGGTGGCATGCAGATTTTTGTCAAGACCTT 2281

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063
              ||||| |||||||||||| | || ||||||  |     |||||||| || ||||| ||||
Db       2282 NACTGGCAAGACCATCACCTTGGAGGTGGAGTCTTCGGACACCATNGACAATGTGAAGGN 2341

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              |||||||| |||||||||| ||||| || |||||||| |  | || ||||| || ||
Db       2342 GAAGATCCAGGACAAGGAAGGAATCCCCCCAGACCAGCAGCGTCTTATTTTTGCTGGCAA 2401

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183
              ||||| || ||||| || |||||  | |||||||||||||||| |||||||| | ||
Db       2402 GCAGCTTGAGGATGGCCGCACCCTAGCAGACTACAACATCCAGAAGGAGTCCACCCTTCA 2461

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243
              |||||||||||  |  | ||||| |||||||||||||| |||||||| || || |||||
Db       2462 CCTGGTGCTCCGCTTNCGCGGTGGTATGCAGATCTTCGTCAAGACCCTCACCGGCAAGAC 2521

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303
              |||||| || |||||||| |    ||||||||| || ||||| ||||| ||||||||| ||
Db       2522 CATCACCCTGGAGGTGGAGTCCTCTGACACCATCGACAATGTGAAGGCGAAGATCCAGGA 2581

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363
              |||||| |||||||| || || |||||| |  | ||||| || || || |||||||| ||
Db       2582 CAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGA 2641

Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423
              ||| |||||||||  | ||||||||||||||||| |||||||| || ||||||||||||||
Db       2642 TGGCCGCACCCTGGCAGACTACAACATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCG 2701

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1462
              |||  | |||||   | |   ||  |  | ||  ||| |
Db       2702 TCTCCGTGGTGGCCAGTAATCCTGGGCCATGAAGCTGTC 2740


RESULT 12
US-09-270-976-5
; Sequence 5, Application US/09270976A
; Patent No. 6706948
; GENERAL INFORMATION:
```

```
;  APPLICANT: Albert, Henrik H.
;  APPLICANT: Wei, Hairong
;  TITLE OF INVENTION: PLANT PROMOTER SEQUENCES AND METHODS OF USE THEREOF
;  FILE REFERENCE: UH-03648
;  CURRENT APPLICATION NUMBER: US/09/270,976A
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 5512
;   TYPE: DNA
;   ORGANISM: Saccharum Hybrid Cultivar H32-8560
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2218)
;   OTHER INFORMATION: The "n" at position 2218 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2227)
;   OTHER INFORMATION: The "n" at position 2227 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2229)
;   OTHER INFORMATION: The "n" at position 2229 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2249)
;   OTHER INFORMATION: The "n" at position 2249 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2282)
;   OTHER INFORMATION: The "n" at position 2282 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2327)
;   OTHER INFORMATION: The "n" at position 2327 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2341)
;   OTHER INFORMATION: The "n" at position 2341 is any nucleotide.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2477)
;   OTHER INFORMATION: The "n" at position 2477 is any nucleotide.
US-09-270-976-5

  Query Match             29.8%;  Score 654.2;  DB 3;  Length 5512;
  Best Local Similarity   80.4%;  Pred. No. 7.9e-180;
  Matches  755;  Conservative    4;  Mismatches  180;  Indels    0;  Gaps    0;

Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583
              |||||||||||||||| |||||||| ||||| |||||||||||||| ||||| ||| |
Db       1802 GATGCAGATCTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCTTC 1861

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643
              ||||: || || :|||||:|||| ||||| || || ||||||||||||||||| |||||
Db       1862 AGACAMTATTGACMATGTCMAGGCTAAGATACAGGACAAGGAAGGCATTCCTCCGGATCA 1921

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703
              ||||||| |||||||||| || || ||||| || ||||| |||||||  :|||||||||
Db       1922 GCAGAGGCTGATCTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGYTGACTACAA 1981

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763
              ||||||||| ||||||||| | |||||||| ||| | ||||| || || ||||||||||
Db       1982 CATCCAGAAGGAGTCCACCSTCCACCTGGTGCTCAGGCTCAGGGGAGGCATGCAAATCTT 2041

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
              ||| ||||| |||||||||||||| ||||| |||||||||||| |   |||||| |||||
Db       2042 CGTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTCGAGTCTTCTGACACGATCGA 2101

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883
              ||||| || || |||||||| |||||||| || || || || ||||||||| |  | ||
Db       2102 CAACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCGGACCAGCAGCGTCTCAT 2161

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943
              ||| || || |||||||| || ||||| |||||||  ||||||||||||||||||  ||
Db       2162 CTTCGCTGGCAAGCAGCTCGAGGATGGCCGCACCCTCGCTGACTACAACATCCAGANGGA 2221

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003
              ||| |   | ||||| |||||| |  | || ||||| |||||||| || || |||||| |
Db       2222 GTCGANTNTCCACCTTGTGCTCAGGTTNAGGGGTGGCATGCAGATTTTTGTCAAGACCTT 2281

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063
              ||||| |||||||||||| | || |||||| |    |||||||| || ||||| ||||
Db       2282 NACTGGCAAGACCATCACCTTGGAGGTGGAGTCTTCGGACACCATNGACAATGTGAAGGN 2341

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              |||||||| |||||||||| ||||| || ||||||||| |   | || ||||| || ||
Db       2342 GAAGATCCAGGACAAGGAAGGAATCCCCCCAGACCAGCAGCGTCTTATTTTTGCTGGCAA 2401

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183
              ||||| || ||||| || ||||| | |||||||||||||||| |||||||| | ||
Db       2402 GCAGCTTGAGGATGGCCGCACCCTAGCAGACTACAACATCCAGAAGGAGTCCACCCTTCA 2461

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243
              |||||||||||   | ||||| |||||||||||||| |||||||| || || |||||
Db       2462 CCTGGTGCTCCGCTTNCGCGGTGGTATGCAGATCTTCGTCAAGACCCTCACCGGCAAGAC 2521

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303
              |||||| || |||||||| |   |||||||| || ||||| ||||| |||||||| ||
Db       2522 CATCACCCTGGAGGTGGAGTCCTCTGACACCATCGACAATGTGAAGGCGAAGATCCAGGA 2581

Qy       1304 CAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1363
              |||||| |||||||| || || |||||| |  | ||||| || || || |||||||| ||
Db       2582 CAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGA 2641

Qy       1364 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1423
              ||| |||||||||| | |||||||||||||||| |||||||| || ||||||||||||
Db       2642 TGGCCGCACCCTGGCAGACTACAACATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCG 2701

Qy       1424 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1462
              |||  | |||||   | |   ||  |  | ||  ||| |
Db       2702 TCTCCGTGGTGGCCAGTAATCCTGGGCCATGAAGCTGTC 2740


RESULT 13
US-08-592-126-140/c
; Sequence 140, Application US/08592126
; Patent No. 5821091
;  GENERAL INFORMATION:
;    APPLICANT:  Gregory Dolganov
;    TITLE OF INVENTION:  Transcripts Encoding Immunomodulatory
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  151
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,126
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sholtz, Charles K.
;      REGISTRATION NUMBER:  38,615
;      REFERENCE/DOCKET NUMBER:  4600-0111
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
```

```
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  795 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  E67.seq
US-08-592-126-140

  Query Match           29.4%;  Score 644.2;  DB 2;  Length 795;
  Best Local Similarity  93.2%;  Pred. No. 2.6e-177;
  Matches  746;  Conservative    0;  Mismatches   42;  Indels   12;  Gaps    7;

Qy        848 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 907
              ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        795 CAAGGAAGGCATTCCTCCTGACCAGGAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 736

Qy        908 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 967
              |||||||||||| |||||||||||||||||||||||||| ||||||||||||||| |||
Db        735 TGGGCGCACCCTNTCTGACTACAACATCCAGAAAGAGTCTNCCCTGCACCTGGTGGTCCN 676

Qy        968 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1027
              ||||||||| || |||||| |||||||||||||||| | ||||||||||||||||||||
Db        675 TCTCAGAGGGGGTATGCAGTTCTTCGTGAAGACCCTNANTGGTAAGACCATCACCCTCGA 616

Qy       1028 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1087
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||| ||
Db        615 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGCTCCAAGACAAGGAAGGGAT 556

Qy       1088 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC-GTACCC 1146
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        555 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGGTGGAAGATGGTCGGTACCC 496

Qy       1147 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG-GTGCTCCGTCTCAGAGGT 1205
              |||||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db        495 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGCGTGCTCCGACTCAGAGGT 436

Qy       1206 GGGATGCAGATCTT-CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              |||||| ||||||| || |||||||||| |||||||||||||||||||||||||||||||
Db        435 GGGATGNAGATCTTCCGNGAAGACCCTGATTGGTAAGACCATCACTCTCGAGGTGGAGCC 376

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              |||||||||||||||| ||||||||||||||||||    | | |    |||| |||
Db        375 GAGTGACACCATTGAG-ATGTCAAGGCAAAGATCCAG----ACAAGGAGCATCCTCTTGA 321

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||| |||| |||| |||||||||| |||||||||||||||||||||||||| |||
Db        320 TCAGCAGAGTTTGACCTTTCCTGGGAAACACCTGGAAGATGGACGCACCCTGTCTGTCTA 261

Qy       1385 CAACATCCAGAA---AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441
              ||||||||||||   ||  |||||||||||||||||||||||||||||||||||||||||
Db        260 CAACATCCAGAAAAGAGTCCCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 201

Qy       1442 GAT-CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||| |||||||||||||||| |||||||||||||||||||||||||||| ||||||||||
Db        200 GATCCTTCGTGAAGACCCTGGCTGGTAAGACCATCACTCTCGAAGTGGAACCGAGTGACA 141

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              |||||||||||||||| |||||| ||||| ||||||||||||||||||||||||||||||
Db        140 CCATTGAGAATGTCAAGNCAAAGACCCAAGNCAAGGAAGGCATCCCTCCTGACCAGCAGA 81

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          80 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 21

Qy       1621 AGAAAGAGTCCACCCTGCAC 1640
              |||||||||||||| || |
Db          20 AGAAAGAGTCCACCCCGCGC 1

RESULT 14
US-09-168-595-140/c
; Sequence 140, Application US/09168595
; Patent No. 6555666
; GENERAL INFORMATION:
;   APPLICANT:  Gregory Dolganov
;   TITLE OF INVENTION:  Transcripts Encoding Immunomodulatory
;   TITLE OF INVENTION:  Polypeptides
;   NUMBER OF SEQUENCES:  151
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Dehlinger & Associates
;     STREET:  350 Cambridge Avenue, Suite 250
;     CITY:  Palo Alto
;     STATE:  CA
;     COUNTRY:  USA
;     ZIP:  94306
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/168,595
;     FILING DATE:
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  08/592,126
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Sholtz, Charles K.
;     REGISTRATION NUMBER:  38,615
;     REFERENCE/DOCKET NUMBER:  4600-0111
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (415) 324-0880
;     TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  140:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  795 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA to mRNA
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   ORIGINAL SOURCE:
;     INDIVIDUAL ISOLATE:  E67.seq
US-09-168-595-140

  Query Match           29.4%;  Score 644.2;  DB 3;  Length 795;
  Best Local Similarity  93.2%;  Pred. No. 2.6e-177;
  Matches  746;  Conservative    0;  Mismatches   42;  Indels   12;  Gaps    7;

Qy        848 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 907
              ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        795 CAAGGAAGGCATTCCTCCTGACCAGGAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 736

Qy        908 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 967
              |||||||||||| |||||||||||||||||||||||||| |||||||||||||| |||
Db        735 TGGGCGCACCCTNTCTGACTACAACATCCAGAAAGAGTCTNCCCTGCACCTGGTGGTCCN 676

Qy        968 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1027
              ||||||||| || |||||| |||||||||||||||| | ||||||||||||||||||||
Db        675 TCTCAGAGGGGGTATGCAGTTCTTCGTGAAGACCCTNANTGGTAAGACCATCACCCTCGA 616

Qy       1028 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1087
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||| ||
Db        615 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGCTCCAAGACAAGGAAGGGAT 556

Qy       1088 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC-GTACCC 1146
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        555 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGGTGGAAGATGGTCGGTACCC 496
```

```
Qy       1147 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG-GTGCTCCGTCTCAGAGGT 1205
             |||||||||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db        495 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGCGTGCTCCGACTCAGAGGT 436

Qy       1206 GGGATGCAGATCTT-CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
             |||||| ||||||| || |||||||||| ||||||||||||||||||||||||||||||||
Db        435 GGGATGNAGATCTTCCGNGAAGACCCTGATTGGTAAGACCATCACTCTCGAGGTGGAGCC 376

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
             |||||||||||||||| |||||||||||||||||     |  | |      |||| |||
Db        375 GAGTGACACCATTGAG-ATGTCAAGGCAAAGATCCAG----ACAAGGAGCATCCTCTTGA 321

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
             ||||||||| |||| |||| |||||||||| |||||||||||||||||||||||||| |||
Db        320 TCAGCAGAGTTTGACCTTTCCTGGGAAACACCTGGAAGATGGACGCACCCTGTCTGTCTA 261

Qy       1385 CAACATCCAGAA---AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441
             ||||||||||||   ||   ||||||||||||||||||||||||||||||||||||||||
Db        260 CAACATCCAGAAAAGAGTCCCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 201

Qy       1442 GAT-CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||| |||||||||||||||| |||||||||||||||||||||||||| ||||||||||
Db        200 GATCCTTCGTGAAGACCCTGGCTGGTAAGACCATCACTCTCGAAGTGGAACCGAGTGACA 141

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             |||||||||||||||| |||||| ||||| ||||||||||||||||||||||||||||||
Db        140 CCATTGAGAATGTCAAGNCAAAGACCCAAGNCAAGGAAGGCATCCCTCCTGACCAGCAGA 81

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          80 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 21

Qy       1621 AGAAAGAGTCCACCCTGCAC 1640
             |||||||||||||||| || |
Db          20 AGAAAGAGTCCACCCCGCGC 1
```

RESULT 15
US-09-404-879A-100
; Sequence 100, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
; LENGTH: 697
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-100

Query Match 28.0%; Score 613.8; DB 3; Length 697;
Best Local Similarity 90.5%; Pred. No. 1.7e-168;
Matches 631; Conservative 45; Mismatches 18; Indels 3; Gaps 3;

```
Qy        192 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA-CAT 250
              ||| |||||||||||||| |||||||||||||||||| |||||||||||||||||| |||
Db          1 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACCAT 60

Qy        251 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 310
              |||||||||||||||||||||||||||||||||||| |||||||||||| ||||||||
Db         61 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGT 120

Qy        311 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 370
              |||||| || ||||| |||||||||| || || |||||||| |||||||||| |||||
Db        121 GAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAA 180

Qy        371 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 430
              :|||||:|||||||||||:|||||||||||||||:||||||||||||||||||||||||
Db        181 YGTCAARGCAAAGATCCARGACAAGGAAGGCATYCCTCCTGACCAGCAGAGGTTGATCTT 240

Qy        431 TGC-CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 489
              ||| :|||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        241 TGCTSGGAAAGCAGCTGGAAGATGGRCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 300

Qy        490 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 549
              |:|||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db        301 CYACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCARATCTTCGTGAAGACCCTGA 360

Qy        550 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 420

Qy        610 AGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAAC 669
              |||||||||||||||||||| |||||||||||||||||||||||||| || ||||
Db        421 AGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAAC 480

Qy        670 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCT-TGCAC 728
              |||||||||||||| || |||||||||||||||||||||||||||||||||| |||||
Db        481 AGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTYTGCAC 540

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
              :||||:||:|||||:||||||:|||:|||||||||::||::||||||||||||:::|||:::
Db        541 YTGGTMCTBCGTCTYAGAGGKGGGRTGCAAATCTWMGTKWAGACACTCACTKKYAAGRYY 600

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              ||||:|::|||::||||:::||:|::||||:||:|:||:||::::|||:||||||:||||
Db        601 ATCAMCMWTGAKKTCGAKYSCASTKWCACTWTCRAKAAMGTYRWWGCAWAGATCCMAGAC 660

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCT 885
              |||||||||||||||||||||||||||||||||||||
Db        661 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCT 697
```

Search completed: March 6, 2006, 19:52:33
Job time : 373 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 19:21:26 ; Search time 8807 Seconds
(without alignments)
11644.974 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccgggga.....aataaagttgttgcattccc 2192

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

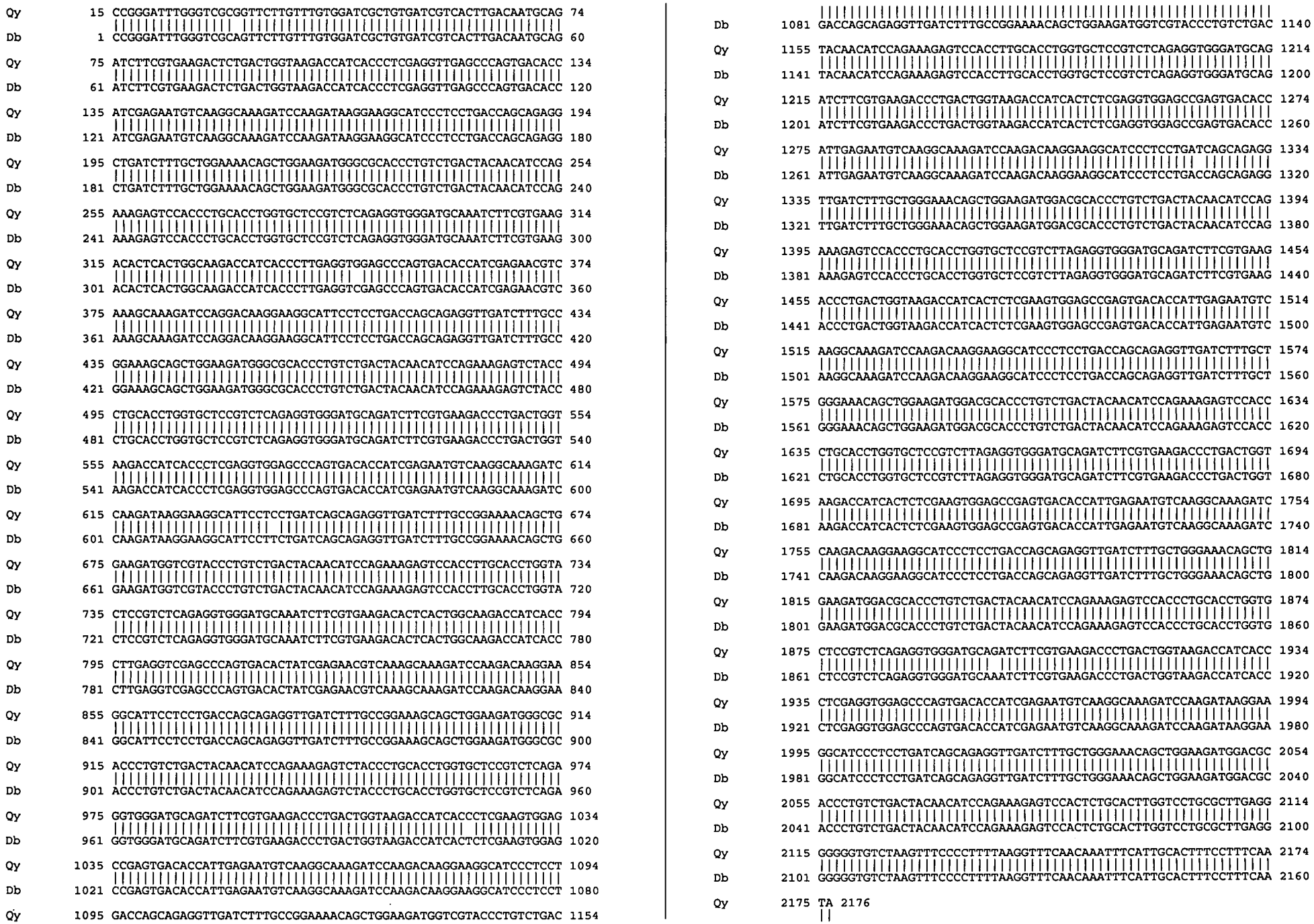
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2152.4	98.2	2162	4	CR621420	CR621420 full-leng
2	1484.8	67.7	1903	4	AK012427	AK012427 Mus muscu
3	1196.2	54.6	1201	4	CR611585	CR611585 full-leng
4	1143.8	52.2	1677	10	CL965182	CL965182 OsIFCC011
5	1132.2	51.7	1698	5	BQ618968	BQ618968 RNOSEQ2A0
6	1132.2	51.7	1698	5	BQ619039	BQ619039 RNOSEQ3A0
7	1132.2	51.7	1698	5	BQ619213	BQ619213 RNOSEQ5A0
8	1030.8	47.0	1098	5	BU902337	BU902337 AGENCOURT
9	1011.4	46.1	1104	1	AL535209	AL535209 AL535209
10	1010.4	46.1	1039	1	AL539566	AL539566 AL539566
11	1001	45.7	1102	1	AL516880	AL516880 AL516880
12	999	45.6	1074	1	AL541647	AL541647 AL541647
13	995.4	45.4	1618	5	BU103666	BU103666 SCCCHR100
14	987.6	45.1	1117	1	AL533345	AL533345 AL533345
15	984.8	44.9	1854	10	CL952190	CL952190 OsJRUA000
16	976.8	44.6	1024	1	AL533035	AL533035 AL533035
17	971.2	44.3	1084	3	BM467727	BM467727 AGENCOURT
18	964.2	44.0	1056	1	AL557176	AL557176 AL557176
19	963.8	44.0	967	4	CR615117	CR615117 full-leng
20	963.6	44.0	1111	1	AL533271	AL533271 AL533271
21	963.6	44.0	1716	3	BP433003	BP433003 BP433003
22	957.8	43.7	973	1	AL535866	AL535866 AL535866
23	953.8	43.5	1087	1	AL514142	AL514142 AL514142
24	952.2	43.4	1117	1	AL540995	AL540995 AL540995
25	951.2	43.4	1060	5	BX461706	BX461706 BX461706
26	950.4	43.4	1050	1	AL534836	AL534836 AL534836
27	949.2	43.3	1114	3	BM455235	BM455235 AGENCOURT
28	948.8	43.3	1118	5	BU902127	BU902127 AGENCOURT
29	947.4	43.2	1092	1	AL538959	AL538959 AL538959
30	946.8	43.2	1086	5	BX417072	BX417072 BX417072
31	946.4	43.2	1057	1	AL559920	AL559920 AL559920
32	944.4	43.1	1325	7	CO635736	CO635736 Contig296
33	943.8	43.1	1085	5	BX440177	BX440177 BX440177
34	943.2	43.0	1086	3	BM541766	BM541766 AGENCOURT
35	937.8	42.8	1085	5	BQ436578	BQ436578 AGENCOURT
36	933.4	42.6	939	1	AL541568	AL541568 AL541568
37	933	42.6	1465	4	AY813424	AY813424 Schistoso
38	931.8	42.5	965	1	AL535900	AL535900 AL535900
39	931.8	42.5	1117	5	BU902021	BU902021 AGENCOURT
40	931.6	42.5	1070	3	BM463129	BM463129 AGENCOURT
41	928.4	42.4	1070	1	AL579176	AL579176 AL579176
42	927.4	42.3	1055	5	BX420118	BX420118 BX420118
43	926.6	42.3	1074	5	BX358407	BX358407 BX358407
44	925.6	42.2	936	7	CR977118	CR977118 CR977118
45	925.4	42.2	1099	3	BM467747	BM467747 AGENCOURT

ALIGNMENTS

RESULT 1
CR621420
LOCUS CR621420 2162 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DE004YK08 of Placenta of Homo sapiens (human).
ACCESSION CR621420
VERSION CR621420.1 GI:50502227
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2162)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2162)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .2162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE004YK08"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"
ORIGIN

Query Match 98.2%; Score 2152.4; DB 4; Length 2162;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 6; Indels 0; Gaps 0;


Qy 15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
Db 1 CCGGGATTTGGGTCGCAGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60
Qy 75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
Db 61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120
Qy 135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
Db 121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180
Qy 195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
Db 181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240
Qy 255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
Db 241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300
Qy 315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
Db 301 ACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTC 360
Qy 375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
Db 361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420
Qy 435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
Db 421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480
Qy 495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
Db 481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540
Qy 555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
Db 541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600
Qy 615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
Db 601 CAAGATAAGGAAGGCATTCCTTCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660
Qy 675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
Db 661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720
Qy 735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
Db 721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780
Qy 795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
Db 781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840
Qy 855 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 914
Db 841 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 900
Qy 915 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 974
Db 901 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 960
Qy 975 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAG 1034
Db 961 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 1020
Qy 1035 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1094
Db 1021 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1080
Qy 1095 GACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 1154
Db 1081 GACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 1140
Qy 1155 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1214
Db 1141 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1200
Qy 1215 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACC 1274
Db 1201 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACC 1260
Qy 1275 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGG 1334
Db 1261 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1320
Qy 1335 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1394
Db 1321 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1380
Qy 1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
Db 1381 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1440
Qy 1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
Db 1441 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1500
Qy 1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
Db 1501 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1560
Qy 1575 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1634
Db 1561 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1620
Qy 1635 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1694
Db 1621 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1680
Qy 1695 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1754
Db 1681 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1740
Qy 1755 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 1814
Db 1741 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 1800
Qy 1815 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1874
Db 1801 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1860
Qy 1875 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1934
Db 1861 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 1920
Qy 1935 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 1994
Db 1921 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 1980
Qy 1995 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 2054
Db 1981 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 2040
Qy 2055 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGG 2114
Db 2041 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGG 2100
Qy 2115 GGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAA 2174
Db 2101 GGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAA 2160
Qy 2175 TA 2176
Db

Db 2161 TA 2162

RESULT 2
AK012427
LOCUS AK012427 1903 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2700054O04 product:POLYUBIQUITIN (UBIQUITIN B) homolog [Mus musculus], full insert sequence.
ACCESSION AK012427
VERSION AK012427.1 GI:12849162
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1903)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.
FEATURES Location/Qualifiers
source 1. .1903
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2700054O04"
/db_xref="taxon:10090"
/clone="2700054O04"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="11 days embryo"
misc_feature 1. .1903
/note="POLYUBIQUITIN (UBIQUITIN B) homolog [Mus musculus] (SPTR|Q63446, evidence: FASTY, 100%ID, 99.6%length, match=912)"
ORIGIN

Query Match 67.7%; Score 1484.8; DB 4; Length 1903;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 1683; Conservative 0; Mismatches 202; Indels 7; Gaps 7;

```
Qy   234 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGT 293
         ||||| ||||||||||||||||||||||||||||||||||||||| || || |||||||||
Db     6 CTGTCAGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGT 65

Qy   294 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCC 353
         || ||||| ||||| |||||||| || || |||||||||||||||||| ||||| ||||||
Db    66 GGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCC 125

Qy   354 AGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGAC 413
         |||||||||| ||||| ||||| |||||||||||||||||||||| ||||| || ||||||
Db   126 AGTGACACCATAGAGAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGAC 185

Qy   414 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 473
         ||||||||| |||||||||| || |||||||||||||||||| |||||||||| ||||||
Db   186 CAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTAC 245

Qy   474 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG-CAGAT 532
         |||||||||||||||| |||||||||||||| || || |||||||||| ||| |||||
Db   246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGTGGCATGCCAGAT 305

Qy   533 CTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCAT 592
         ||| |||||||||||||| || |||||||||||||| ||||| ||||||||||||||||||
Db   306 CTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCAT 365

Qy   593 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGA-AGGCATTCCTCCTGATCAGCAGAGGT 651
         |||||||||||||||||||||||||| ||||| ||||| || ||||| |||||||||||
Db   366 AGAGAATGTCAAGGCAAAGATCCAGGACAAGGAGGGGCATCCCCCCTGACCAGCAGAGGC 425

Qy   652 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 711
         |||||||||| || || |||||||||||||| || |||||||| ||||||||||||||||
Db   426 TGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGA 485

Qy   712 AAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 771
         |||||||||| |||||||||| || || |||||||||| ||||| ||||| |||||||
Db   486 AAGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGTGGCATGCAGATCTTTGTGAAGA 545

Qy   772 CACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCA 831
         | || || |||||||||||||||||| |||||||||||||||||||| || ||||| ||||
```

```
Db        546 CCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATAGAGAATGTCA 605

Qy        832 AAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 891

Db        606 AGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAG 665

Qy        892 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 951

Db        666 GCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAAGAGTCCACCC 725

Qy        952 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1011

Db        726 TGCACCTGGTCCTTCGCCTCAGAGGTGGCATGCAGATCTTTGTGAAGACCCTGACAGGCA 785

Qy       1012 AGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1071

Db        786 AGACCATCACCCTGGAGGTCGAGCCCAGTGACACCATAGAGAATGTCAAGGCAAAGATCC 845

Qy       1072 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 1131

Db        846 AGGACAAGGAGGGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGG 905

Qy       1132 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGC 1191

Db        906 AAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTCC 965

Qy       1192 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1251

Db        966 TTCGCCTCAGAGGTGGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCC 1025

Qy       1252 TCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA-AGACAAGGAA 1310

Db       1026 TGGAGGTCGAGCCCAGTGACACCATAGAGAATGTCAAGGCAAAGATCCAGGGACAAGGAG 1085

Qy       1311 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1370

Db       1086 GGCATCCCCCCTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGC 1145

Qy       1371 ACCCTGTCTGACTACAACATCCAG-AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAG 1429

Db       1146 ACCCTGTCAGACTACAACATCCAGAAAAGAGTCCACCCTGCACCTGGTCCTTCGCCTCAG 1205

Qy       1430 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGA 1489

Db       1206 AGGTGGCATGCAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGA 1265

Qy       1490 GCCGAGTGACACCATTGAG-AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1548

Db       1266 GCCCAGTGACACCATAGAGAAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCC 1325

Qy       1549 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1608

Db       1326 CTGACCAGCAGAGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAG 1385

Qy       1609 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATG- 1667

Db       1386 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGTGGCATGC 1445

Qy       1668 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGAC 1727

Db       1446 CAGATCTTTGTGAAGACCCTGACAGGCAAGACCATCACCCTGGAGGTCGAGCCCAGTGAC 1505

Qy       1728 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 1787

Db       1506 ACCATAGAGAATGTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCCCCCTGACCAGCAG 1565

Qy       1788 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1847

Db       1566 AGGCTGATCTTTGCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATC 1625

Qy       1848 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 1907

Db       1626 CAGAAAGAGTCCACCCTGCACCTGGTCCTTCGCCTCAGAGGTGGCATGCAGATCTTTGTG 1685

Qy       1908 AAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAAT 1967

Db       1686 AAGACCCTGACAGGCAAGACCATCACCCTGGACGTCGAGCCCAGTGTTACCACCAAGAAG 1745

Qy       1968 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 2027

Db       1746 GTCAAGGCAAAGATCCAGGACAAGGAGGGCATCCC-CCTGACCAGCAGAGGCTGATCTTT 1804

Qy       2028 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 2087

Db       1805 GCAGGCAAGCAGCTGGAAGATGGCCGCACCCTGTCAGACTACAACATCCAGAAAGAGTCC 1864

Qy       2088 ACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119

Db       1865 ACCGTGCACTGGTCCTTCGCCTCAGAGGTGGG 1896
```

RESULT 3
CR611585

```
LOCUS       CR611585               1201 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CL0BB015ZB04 of Neuroblastoma of Homo
            sapiens (human).
ACCESSION   CR611585
VERSION     CR611585.1  GI:50492392
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CL0BB015ZB04"
                     /tissue_type="Neuroblastoma"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             54.6%;  Score 1196.2;  DB 4;  Length 1201;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1198;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy         991 TCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTG 1050

Db           1 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTG 60

Qy        1051 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGA 1110

Db          61 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGA 120

Qy        1111 TCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAG 1170
```

```
Db       121 TCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAG 180

Qy      1171 AGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCC 1230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCC 240

Qy      1231 TGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGG 1290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGG 300

Qy      1291 CAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGA 1350
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       301 CAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGA 360

Qy      1351 AACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 1410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 420

Qy      1411 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 480

Qy      1471 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1530
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 540

Qy      1531 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 1590
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 600

Qy      1591 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 1650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 660

Qy      1651 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 1710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 720

Qy      1711 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 1770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 780

Qy      1771 TCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCC 1830
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCC 840

Qy      1831 TGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 1890
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 900

Qy      1891 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 1950
             ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 960

Qy      1951 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATC 2010
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATC 1020

Qy      2011 AGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACA 2070
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACA 1080

Qy      2071 ACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTT 2130
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTT 1140

Qy      2131 CCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 1200

Qy      2191 C 2191
             |
Db       1201 C 1201
```

```
RESULT 4
CL965182
LOCUS       CL965182                1677 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC011870 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL965182
VERSION     CL965182.1  GI:52385056
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1677)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .1677
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

  Query Match             52.2%;  Score 1143.8;  DB 10;  Length 1677;
  Best Local Similarity   82.3%;  Pred. No. 1.4e-301;
  Matches 1313;  Conservative    0;  Mismatches  282;  Indels    0;  Gaps    0;

Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
              |||||||||| |||||||| || || || ||||||||||||||||||||||| ||
Db          1 ATGCAGATCTTTGTGAAGACCCTCACCGGCAAGACCATCACCCTCGAGGTTGAGTCCTCG 60

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
              |||||||| || |||||||||| |||||||| || ||||| |||||||| || ||||||
Db         61 GACACCATTGACAATGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAG 120

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
              ||| | || ||||| ||||| || ||||| || ||||| ||||||||| | |||||||||
Db        121 CAGCGTCTCATCTTCGCTGGCAAGCAGCTTGAGGATGGCCGCACCCTGGCCGACTACAAC 180

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
              ||||||| |||||||||| ||||| |||||| | ||||| || || ||||| ||||||
Db        181 ATCCAGAAGGAGTCCACCCTCCACCTTGTGCTCAGGCTCAGGGGAGGCATGCAGATCTTC 240

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
              || |||||  | |||||||||||||||||||||||||||| ||| |   ||||||||| ||
Db        241 GTCAAGACCTTGACTGGCAAGACCATCACCCTTGAGGTCGAGTCGTCTGACACCATTGAC 300

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
              || ||||| || |||||||||||||||||| ||||| || || ||||||||| |  | |||
Db        301 AATGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATC 360

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
              || || || |||||||| || ||||| ||||||||| | |||||||||||||||| |||
Db        361 TTCGCTGGCAAGCAGCTTGAGGATGGCCGCACCCTGGCCGACTACAACATCCAGAAGGAG 420

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
```

```
Db        421 TCTACCCTCCACCTTGTGCTCAGGCTCAGGGGAGGCATGCAGATCTTCGTCAAGACCTTG 480

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608

Db        481 ACTGGCAAGACCATCACCCTTGAGGTCGAGTCGTCTGACACCATTGACAATGTCAAGGCC 540

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668

Db        541 AAGATCCAGGACAAGGAGGGCATCCCCCCAGACCAGCAGCGTCTCATCTTCGCCGGCAAG 600

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728

Db        601 CAGCTGGAGGATGGCCGCACCCTTGCTGACTACAACATCCAGAAGGAGTCCACCCTCCAC 660

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788

Db        661 CTTGTGCTCAGGCTCAGGGGAGGTATGCAGATCTTCGTCAAGACCCTGACCGGCAAGACC 720

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848

Db        721 ATCACCCTCGAGGTCGAGTCCTCGGACACGATCGACAACGTGAAGGCCAAGATCCAGGAC 780

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908

Db        781 AAGGAGGGCATCCCCCCGGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGAC 840

Qy        909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968

Db        841 GGCCGCACCCTTGCCGACTACAACATCCAGAAGGAATCCACCCTCCACCTTGTGCTTAGG 900

Qy        969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028

Db        901 CTCAGGGGAGGTATGCAGATCTTCGTCAAGACCCTGACTGGCAAGACCATCACACTTGAG 960

Qy       1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088

Db        961 GTCGAGTCCTCGGACACGATCGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATC 1020

Qy       1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148

Db       1021 CCCCCGGACCAGCAGCGTCTCATCTTCGCTGGCAAGCAGCTGGAGGACGGCCGCACCCTT 1080

Qy       1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208

Db       1081 GCCGACTACAACATCCAGAAGGAGTCCACCCTCCACCTTGTGCTCAGGCTCAGGGGTGGT 1140

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268

Db       1141 ATGCAGATCTTCGTCAAGACCCTGACCGGCAAGACCATCACGCTTGAGGTCGAGTCCTCG 1200

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328

Db       1201 GACACGATCGACAACGTGAAGGCCAAGATCCAGGACAAGGAGGGTATCCCCCCGGACCAG 1260

Qy       1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388

Db       1261 CAGCGCCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCACCTTGGCTGACTACAAC 1320

Qy       1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448

Db       1321 ATCCAGAAGGAGTCCACCCTTCACCTGGTTCTCAGGCTCAGGGGTGGGATGCAGATCTTC 1380

Qy       1449 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1508

Db       1381 GTGAAGACCCTGACTGGCAAGACCATTACCCTTGAGGTTGAGTCGTCCGACACTATTGAC 1440

Qy       1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568

Db       1441 AACGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAGCAGCGTCTGATC 1500

Qy       1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628

Db       1501 TTTGCTGGTAAGCAGCTTGAGGATGGCCGCACCCTTGCGGATTACAACATCCAGAAGGAG 1560

Qy       1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1663

Db       1561 TCCACACTCCACCTGGTGCTCCGCCTCCGTGGTGG 1595
```

RESULT 5
BQ618968

```
LOCUS       BQ618968               1698 bp    mRNA    linear   EST 06-SEP-2002
DEFINITION  RNOSEQ2A09_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
            mays cDNA clone RNOSEQ2A09_SK.ab1 similar to polyubiquitin [Pinus
            sylvestris], mRNA sequence.
ACCESSION   BQ618968
VERSION     BQ618968.1  GI:21620962
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1698)
  AUTHORS   Wang,H. and Bohnert,H.J.
  TITLE     Genomics of plant stress tolerance
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Mark Fredricksen
            Department of Plant Biology
            University of Illinois
            1201 W. Gregory Dr., Urbana, IL 61801, USA
            Tel: 2172655473
            Email: bohnertlab@life.uiuc.edu.
FEATURES             Location/Qualifiers
     source          1. .1698
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="RNOSEQ2A09_SK.ab1"
                     /tissue_type="Roots"
                     /dev_stage="2 weeks old"
                     /clone_lib="Salt stressed Zea mays roots cDNA library"
                     /note="Vector: pBluescript SK+; Stressed 24 hours at 150
                     mM NaCl"
ORIGIN
```

Query Match 51.7%; Score 1132.2; DB 5; Length 1698;
Best Local Similarity 81.6%; Pred. No. 2.2e-298;
Matches 1322; Conservative 0; Mismatches 298; Indels 1; Gaps 1;

```
Qy         48 CGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATC 107

Db         53 CGTCGGCACCTCCGCTTCAAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATC 112

Qy        108 ACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 167

Db        113 ACCCTCGAGGTGGAGTCGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAG 172

Qy        168 GAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGG 227

Db        173 GAGGGCATCCCCCCAGACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGG 232

Qy        228 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 287

Db        233 CGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTC 292

Qy        288 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTG 347

Db        293 AGGGGAGGCATGCAGATCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTG 352

Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407

Db        353 GAGTCCTCTGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCT 412
```

```
Qy       408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
            || ||||||||| || | |||||||| || |||||||| || || |||||||| ||   |
Db       413 CCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCC 472

Qy       468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
            |||||||||||||||| |||    ||||| ||| ||||||| || ||||| || || |||
Db       473 GACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATG 532

Qy       528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
            |||||||||||||||||||||| || ||||| ||||||||||||||||| |     |||
Db       533 CAGATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGAC 592

Qy       588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
            |||||||| ||||| ||||| |||||||| || ||||| |||||||| || || ||||||
Db       593 ACCATCGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAG 652

Qy       648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
            | |||||||||| ||| | |||||||| ||||| || |||||  | || |||||||||
Db       653 CGTCTGATCTTTG-CGGGCAGCAGCTGGAGGATGGCCGCACCCTTGCGGATTACAACATC 711

Qy       708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
            ||||| |||  |||| | |||||||| |||||||||| ||||| ||||| ||||| |||
Db       712 CAGAAGGAGAGCACCCTCCACCTGGTGCTCCGTCTCAGGGGTGGTATGCAGATCTTTGTG 771

Qy       768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
            |||||||||||||||||| |||||||||||| ||| |    || || || || ||
Db       772 AAGACACTCACTGGCAAGACAATCACCCTTGAGGTGGAGTCTTCGGATACCATTGACAAT 831

Qy       828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
            ||||| || |||||||| |||||||| ||||| || || ||||||||| |   | |||||
Db       832 GTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCACCCGACCAGCAGCGCCTCATCTTC 891

Qy       888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
            ||||| |||||||||| ||||| ||||||||| | || ||||||||||||||| |||
Db       892 GCCGGCAAGCAGCTGGAGGATGGCCGCACCCTGGCGGATTACAACATCCAGAAGGAGAGC 951

Qy       948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
            || || |||||||||||||| ||||| ||||| |||||||| || ||||||||  |||||
Db       952 ACTCTCCACCTGGTGCTCCGCCTCAGGGGTGGCATGCAGATTTTTGTGAAGACATTGACT 1011

Qy      1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
            || |||||||||||| | || ||||||      |||||||||||| ||||| ||||| |||
Db      1012 GGCAAGACCATCACCTTGGAGGTGGAGAGCTCTGACACCATTGACAATGTGAAGGCCAAG 1071

Qy      1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
            ||||| ||||||| ||||| || || ||||||||| | ||||||||| || || |||
Db      1072 ATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGTCTGATCTTTGCGGGCAAGCAG 1131

Qy      1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
            ||||| ||||| || || ||  | |||||||||||||||| |||  |||| | |||||
Db      1132 CTGGAGGATGGCCGCACTCTCGCGGACTACAACATCCAGAAGGAGAGCACCCTTCACCTT 1191

Qy      1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
            || ||||| |||||||||| |||||||||| || |||||||||||||| || |||||
Db      1192 GTTCTCCGCCTCAGAGGTGGTATGCAGATCTTTGTAAAGACCCTGACTGGAAAAACCATA 1251

Qy      1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
            || ||||  |||      |||||||| || |||||||| |||||||| ||||||
Db      1252 ACCCTGGAGGGTGAGAGCTCGGACACCATCGACAACGTCAAGGCCAAGATCCAGGACAAG 1311

Qy      1308 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1367
            || |||||||||| || |||||| || | ||||||||||||||| ||||| || || ||
Db      1312 GAGGGCATCCCTCCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACGGG 1371

Qy      1368 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1427
            ||||| ||  |||||||||||||||||| |||  |||||| ||| ||||||| || ||
Db      1372 CGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTC 1431

Qy      1428 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1487
            || || || |||||||||||||||||||||||||| || ||||| ||||| ||||| |||
Db      1432 AGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCTCGAGGTG 1491

Qy      1488 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1547
```

```
            ||| |    |||||||||| || |||||||| |||||||| |||||||| ||||| ||
Db      1492 GAGTCTTCAGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATTCCC 1551

Qy      1548 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 1607
            || ||||||||| || | |||||||||| || ||||| || || || ||||| ||  |
Db      1552 CCAGACCAGCAGCGGCTCATCTTTGCTGGAAAGCAGCTTGAGGACGGGCGCACGCTTGCC 1611

Qy      1608 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATG 1667
            |||||||||||||||||| |||  |||||| ||| ||||||| || || || || || |||
Db      1612 GACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATG 1671

Qy      1668 C 1668
            |
Db      1672 C 1672
```

RESULT 6
BQ619039
LOCUS BQ619039 1698 bp mRNA linear EST 06-SEP-2002
DEFINITION RNOSEQ3A06_SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RNOSEQ3A06_SK.ab1 similar to polyubiquitin [Pinus sylvestris], mRNA sequence.
ACCESSION BQ619039
VERSION BQ619039.1 GI:21621033
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1698)
AUTHORS Wang,H. and Bohnert,H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
FEATURES Location/Qualifiers
source 1. .1698
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="RNOSEQ3A06_SK.ab1"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
ORIGIN

Query Match 51.7%; Score 1132.2; DB 5; Length 1698;
Best Local Similarity 81.6%; Pred. No. 2.2e-298;
Matches 1322; Conservative 0; Mismatches 298; Indels 1; Gaps 1;

```
Qy        48 CGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATC 107
            ||  |  | |  | ||| |  |||||||||| ||||| || |||||||| ||||| |||
Db        53 CGTCGGCACCTCCGCTTCAAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATC 112

Qy       108 ACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 167
            |||||||||| ||| |    ||||||||| || || || ||||| |||||||| || |||
Db       113 ACCCTCGAGGTGGAGTCGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAG 172

Qy       168 GAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGG 227
            || |||||||| || |||||||| |||| |||||||||| |||||||| || || |||
Db       173 GAGGGCATCCCCCCAGACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGG 232

Qy       228 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 287
            ||||| ||  |||||||||||||||||| |||  |||||| ||||| ||||||||||||
Db       233 CGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTC 292
```

```
Qy        288 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTG 347
              || || || ||||| ||||| ||||| || || || |||||||| |||||||| ||||||
Db        293 AGGGGAGGCATGCAGATCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTG 352

Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
              ||| ||  |||||||| || |||||||| || ||||||||||||||| ||||| |||
Db        353 GAGTCCTCTGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCT 412

Qy        408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
              || |||||||| || | |||||||| || |||||||| || || |||||||| ||   |
Db        413 CCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCC 472

Qy        468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
              |||||||||||||||||  |||    ||||| ||| ||||||| || ||||| || || |||
Db        473 GACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATG 532

Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
              |||||||||||||||||||||| || ||||| |||||||||||||||||  |    |||
Db        533 CAGATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGAC 592

Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
              |||||||| ||||| ||||| |||||||| || ||||| |||||||| || || ||||||
Db        593 ACCATCGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAG 652

Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
              |  ||||||||| |||  | |||||||| ||||| || |||||  | || |||||||||
Db        653 CGTCTGATCTTTG-CGGGCAGCAGCTGGAGGATGGCCGCACCCTTGCGGATTACAACATC 711

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              ||||| |||  |||| | |||||||| |||||||||| ||||| ||||| ||||| |||
Db        712 CAGAAGGAGAGCACCCTCCACCTGGTGCTCCGTCTCAGGGGTGGTATGCAGATCTTTGTG 771

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
              |||||||||||||||||| |||||||||||||| ||| |    || || || || ||
Db        772 AAGACACTCACTGGCAAGACAATCACCCTTGAGGTGGAGTCTTCGGATACCATTGACAAT 831

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||| || |||||||| |||||||| ||||| || || ||||||||| |  | |||||
Db        832 GTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCACCCGACCAGCAGCGCCTCATCTTC 891

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||| |||||||||| ||||| ||||||||  | || |||||||||||||| |||
Db        892 GCCGGCAAGCAGCTGGAGGATGGCCGCACCCTGGCGGATTACAACATCCAGAAGGAGAGC 951

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              || || |||||||||||||| ||||| ||||| |||||||| || ||||||||  |||||
Db        952 ACTCTCCACCTGGTGCTCCGCCTCAGGGGTGGCATGCAGATTTTTGTGAAGACATTGACT 1011

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              || |||||||||||| | || ||||||      |||||||||||| ||||| ||||| |||
Db       1012 GGCAAGACCATCACCTTGGAGGTGGAGAGCTCTGACACCATTGACAATGTGAAGGCCAAG 1071

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
              ||||| |||||||| ||||| || || ||||||||| |  |||||||||| || || |||
Db       1072 ATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGTCTGATCTTTGCGGGCAAGCAG 1131

Qy       1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
              ||||| ||||| || || ||  | |||||||||||||||| |||  |||| | ||||||
Db       1132 CTGGAGGATGGCCGCACTCTCGCGGACTACAACATCCAGAAGGAGAGCACCCTTCACCTT 1191

Qy       1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
              || ||||| |||||||||| |||||||||| || |||||||||||||| || ||||||
Db       1192 GTTCTCCGCCTCAGAGGTGGTATGCAGATCTTTGTAAAGACCCTGACTGGAAAAACCATA 1251

Qy       1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
              || || ||||  |||       |||||||| || || |||||||| |||||||| ||||||
Db       1252 ACCCTGGAGGGTGAGAGCTCGGACACCATCGACAACGTCAAGGCCAAGATCCAGGACAAG 1311

Qy       1308 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1367
              || |||||||||| || |||||| || | |||||||||||||| ||||| || || ||
Db       1312 GAGGGCATCCCTCCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACGGG 1371

Qy       1368 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1427
              ||||| ||  |||||||||||||||||| |||  |||||| ||| ||||||| || ||
Db       1372 CGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTC 1431

Qy       1428 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1487
              || || || |||||||||||||||||||||||||| || ||||| ||||| ||||| |||
Db       1432 AGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCTCGAGGTG 1491

Qy       1488 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1547
              ||| |    ||||||||||| || |||||||| |||||||| |||||||| ||||| ||
Db       1492 GAGTCTTCAGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATTCCC 1551

Qy       1548 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 1607
              || |||||||||  || | |||||||||||  || ||||| || || || ||||| ||   |
Db       1552 CCAGACCAGCAGCGGCTCATCTTTGCTGGAAAGCAGCTTGAGGACGGGCGCACGCTTGCC 1611

Qy       1608 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATG 1667
              |||||||||||||||||  |||  |||||| ||| ||||||| || || || || || |||
Db       1612 GACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATG 1671

Qy       1668 C 1668
              |
Db       1672 C 1672
```

RESULT 7
BQ619213
LOCUS BQ619213 1698 bp mRNA linear EST 06-SEP-2002
DEFINITION RNOSEQ5A09_SK.ab1 Salt stressed Zea mays roots cDNA library Zea mays cDNA clone RNOSEQ5A09_SK.ab1 similar to polyubiquitin [Pinus sylvestris], mRNA sequence.
ACCESSION BQ619213
VERSION BQ619213.1 GI:21621207
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1698)
AUTHORS Wang,H. and Bohnert,H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished (2002)
COMMENT Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.
FEATURES Location/Qualifiers
source 1. .1698
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="RNOSEQ5A09_SK.ab1"
/tissue_type="Roots"
/dev_stage="2 weeks old"
/clone_lib="Salt stressed Zea mays roots cDNA library"
/note="Vector: pBluescript SK+; Stressed 24 hours at 150 mM NaCl"
ORIGIN

Query Match 51.7%; Score 1132.2; DB 5; Length 1698;
Best Local Similarity 81.6%; Pred. No. 2.2e-298;
Matches 1322; Conservative 0; Mismatches 298; Indels 1; Gaps 1;

```
Qy         48 CGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATC 107
              ||  |  | |  | ||| |  |||||||||||| ||||| || |||||||| ||||| |||
Db         53 CGTCGGCACCTCCGCTTCAAGATGCAGATCTTTGTGAAAACCCTGACTGGCAAGACTATC 112

Qy        108 ACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 167
              |||||||||||| ||| |   ||||||||| || || || ||||| |||||||| || |||
```

```
Db        113 ACCCTCGAGGTGGAGTCGTCTGACACCATTGACAACGTTAAGGCCAAGATCCAGGACAAG 172
Qy        168 GAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGG 227
Db        173 GAGGGCATCCCCCCAGACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGGACGGG 232
Qy        228 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 287
Db        233 CGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACCTTGTGCTCCGTCTC 292
Qy        288 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTG 347
Db        293 AGGGGAGGCATGCAGATCTTTGTGAAAACCCTGACCGGCAAGACTATCACCCTCGAGGTG 352
Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
Db        353 GAGTCCTCTGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCT 412
Qy        408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
Db        413 CCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACGGGCGCACGCTTGCC 472
Qy        468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
Db        473 GACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATG 532
Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
Db        533 CAGATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCTCGAGGTGGAGTCTTCAGAC 592
Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
Db        593 ACCATCGACAATGTGAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAG 652
Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
Db        653 CGTCTGATCTTTG-CGGGCAGCAGCTGGAGGATGGCCGCACCCTTGCGGATTACAACATC 711
Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
Db        712 CAGAAGGAGAGCACCCTCCACCTGGTGCTCCGTCTCAGGGGTGGTATGCAGATCTTTGTG 771
Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
Db        772 AAGACACTCACTGGCAAGACAATCACCCTTGAGGTGGAGTCTTCGGATACCATTGACAAT 831
Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
Db        832 GTCAAGGCCAAGATCCAGGACAAGGAGGGCATCCCACCCGACCAGCAGCGCCTCATCTTC 891
Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
Db        892 GCCGGCAAGCAGCTGGAGGATGGCCGCACCCTGGCGGATTACAACATCCAGAAGGAGAGC 951
Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
Db        952 ACTCTCCACCTGGTGCTCCGCCTCAGGGGTGGCATGCAGATTTTTGTGAAGACATTGACT 1011
Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
Db       1012 GGCAAGACCATCACCTTGGAGGTGGAGAGCTCTGACACCATTGACAATGTGAAGGCCAAG 1071
Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
Db       1072 ATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGTCTGATCTTTGCGGGCAAGCAG 1131
Qy       1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
Db       1132 CTGGAGGATGGCCGCACTCTCGCGGACTACAACATCCAGAAGGAGAGCACCCTTCACCTT 1191
Qy       1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
Db       1192 GTTCTCCGCCTCAGAGGTGGTATGCAGATCTTTGTAAAGACCCTGACTGGAAAAACCATA 1251
Qy       1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
Db       1252 ACCCTGGAGGGTGAGAGCTCGGACACCATCGACAACGTCAAGGCCAAGATCCAGGACAAG 1311
Qy       1308 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1367
Db       1312 GAGGGCATCCCTCCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACGGG 1371
Qy       1368 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1427
Db       1372 CGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTC 1431
Qy       1428 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1487
Db       1432 AGGGGAGGCATGCAGATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCTCGAGGTG 1491
Qy       1488 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1547
Db       1492 GAGTCTTCAGACACCATTGACAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATTCCC 1551
Qy       1548 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 1607
Db       1552 CCAGACCAGCAGCGGCTCATCTTTGCTGGAAAGCAGCTTGAGGACGGGCGCACGCTTGCC 1611
Qy       1608 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATG 1667
Db       1612 GACTACAACATCCAGAAGGAGAGCACCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATG 1671
Qy       1668 C 1668
Db       1672 C 1672
```

RESULT 8
BU902337
LOCUS BU902337 1098 bp mRNA linear EST 17-OCT-2002
DEFINITION AGENCOURT_10127301 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6503328 5', mRNA sequence.
ACCESSION BU902337
VERSION BU902337.1 GI:24084250
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1098)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14060 row: n column: 01
High quality sequence stop: 810.
FEATURES Location/Qualifiers
source 1. .1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6503328"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

```
                    Average insert size 2.1 kb. "
ORIGIN

  Query Match             47.0%;  Score 1030.8;  DB 5;  Length 1098;
  Best Local Similarity   97.3%;  Pred. No. 1.1e-270;
  Matches 1069;  Conservative    0;  Mismatches   28;  Indels    2;  Gaps    2;

Qy       1018 TCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 1077
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 60

Qy       1078 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 1137
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 120

Qy       1138 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTC 1197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTC 180

Qy       1198 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGG 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGG 240

Qy       1258 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 300

Qy       1318 CTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 1377
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 360

Qy       1378 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 420

Qy       1438 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTG 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTG 480

Qy       1498 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 540

Qy       1558 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 600

Qy       1618 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 660

Qy       1678 TGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGA 1737
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        661 TGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGA 720

Qy       1738 ATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCT 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCT 780

Qy       1798 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 1857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 840

Qy       1858 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 1917
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        841 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGA 900

Qy       1918 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 1977
             ||||  ||| |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        901 CTGGNTAGA-CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAAAATGTCAAGGCAA 959

Qy       1978 AGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAAC 2037
             |||||||||||| | | ||   |||||||||||  | |||||||||||||||||||||||
```

```
Db        960 AGATCCAAGATAGGAAGGGATCCCCTCCTGATCACAAAAGGTTGATCTTTGCTGGGAAAC 1019

Qy       2038 AGCTGGAAGATGG-ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCAC 2096
             |||||||||||| || |||||| | ||||||||||||||||||| |  || |||||| |
Db       1020 CGCTGGAAGATGGAACCCACCCTTTTTGACTACAACATCCAGAAAAAAGCCCCTCTGCCC 1079

Qy       2097 TTGGTCCTGCGCTTGAGGG 2115
             |||| ||||| ||||| ||
Db       1080 TTGGCCCTGCCCTTGAAGG 1098


RESULT 9
AL535209
LOCUS       AL535209                1104 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535209 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF008YB24 5-PRIME, mRNA sequence.
ACCESSION   AL535209
VERSION     AL535209.3  GI:45711087
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1104)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2Q01 this sequence version replaced gi:30541755.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF008DA12QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1104
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF008YB24"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             46.1%;  Score 1011.4;  DB 1;  Length 1104;
  Best Local Similarity   95.9%;  Pred. No. 2.2e-265;
  Matches 1064;  Conservative   22;  Mismatches   18;  Indels    5;  Gaps    5;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCAGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180
```

```
Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 900

Qy        915 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 974
              :||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 -MCCYGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 959

Qy        975 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAG 1034
              |||||||||||||:||||||:||:| |||||:|||||| | |||||| ||||||||||:|
Db        960 GGTGGGATGCAGAWCTTCGTKAARA-CCTGAMTGGTAA-AACATCACTCTCGAAGTGGRG 1017

Qy       1035 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1094
              |||||||||||:||||||||||:|| |::||||: | |    :|:||| ||||||:||:|
Db       1018 -CGAGTGACACCMTTGAGAATGTYAAAGSMAAGAWTCCAATAMARGAA-GCATCCYTCYT 1075

Qy       1095 GACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              :||:||:| || |   | ||: : |:|||
Db       1076 KACMAGMAAAGTTGATTTTTYCSGGRAAA 1104


RESULT 10
AL539566
LOCUS       AL539566               1039 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL539566 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF036YO05 5-PRIME, mRNA sequence.
ACCESSION   AL539566
VERSION     AL539566.3  GI:45715224
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1039)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31264133.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF036AH03QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1039
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF036YO05"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             46.1%;  Score 1010.4;  DB 1;  Length 1039;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-265;
  Matches 1016;  Conservative    9;  Mismatches    0;  Indels    1;  Gaps    1;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 67

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 127

Qy        140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 187

Qy        200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 247

Qy        260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 307

Qy        320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 367

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 427

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 487
```

```
Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 547

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 607

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 667

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 727

Qy        740 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 787

Qy        800 GGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCAT 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 GGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCAT 847

Qy        860 TCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCT 919
              ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        848 TCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGRAAGCAGCTGGAAGATGGGCGCACCCT 907

Qy        920 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 967

Qy        980 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 1039
              |||||||||:|||:|||||||||||||||||||||||:::::| ||||||||||||||:||||
Db        968 GATGCAGATYTTCKTGAAGACCCTGACTGGTAAGACMWYMA-CCTCGAAGTGGAGSCGAG 1026

Qy       1040 TGACAC 1045
              ||:|||
Db       1027 TGRCAC 1032
```

```
RESULT 11
AL516880
LOCUS       AL516880               1102 bp    mRNA    linear   EST 23-MAR-2004
DEFINITION  AL516880 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA007YA09 5-PRIME, mRNA sequence.
ACCESSION   AL516880
VERSION     AL516880.3  GI:45653489
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1102)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30491945.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DA007AA05QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1102
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA007YA09"
                     /tissue_type="NEUROBLASTOMA"
                     /clone_lib="Homo sapiens NEUROBLASTOMA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN
```

```
  Query Match            45.7%;  Score 1001;  DB 1;  Length 1102;
  Best Local Similarity  96.6%;  Pred. No. 1.5e-262;
  Matches 1008;  Conservative   21;  Mismatches   13;  Indels    1;  Gaps    1;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              || |||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         13 GGATCGCAGTTCTTGTTTGTGGATTGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 72

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 132

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 192

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 252

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 312

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        313 GGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 372

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 432

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 492

Qy        504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 552

Qy        564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 612

Qy        624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 672

Qy        684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        673 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCYTGCACCTGGTACTCCGTCTC 732

Qy        744 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 792

Qy        804 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 852
```

```
Qy        864 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 923
              ||:||||:|||||||||||||||||||||||:||:||||:|||||||||||||||||||||
Db        853 CCYGACCMGCAGAGGTTGATCTTTGCCGGRAARCAGCHGGAAGATGGGCGCACCCTGTCT 912

Qy        924 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 983
              ||:||||||||||:|||||:|||||||:|||||||||||||||||||:|||||||::|
Db        913 GAMTACAACATCCARAAAGARTCTACCCYGCACCTGGTGCTCCGTCTCARAGGTGGGRKG 972

Qy        984 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 1043
              ||||||||||:|||||||:||| | :|:| |||:|| |||||  |||||| ||||| |
Db        973 CAGATCTTCGTGWAGACCCTSACTTGGWARAACATMACTCTCGAGGTGGAG-CGAGTTAA 1031

Qy       1044 ACCATTGAGAATGTCAAGGCAAA 1066
              ||:|||||||||:| |||: |||
Db       1032 ACMATTGAGAATKTAAAGVAAAA 1054
```

```
RESULT 12
AL541647
LOCUS       AL541647               1074 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL541647 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE003YJ22
            5-PRIME, mRNA sequence.
ACCESSION   AL541647
VERSION     AL541647.3  GI:45717228
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1074)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:30546021.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DE003DE11QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1074
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE003YJ22"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN
```

```
  Query Match             45.6%;  Score 999;  DB 1;  Length 1074;
  Best Local Similarity   94.6%;  Pred. No. 5.4e-262;
  Matches 1015;  Conservative   33;  Mismatches   23;  Indels    2;  Gaps    2;

Qy         14 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 73
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCGGGATTTGGGTCGCAGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 60

Qy         74 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 120

Qy        134 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 180

Qy        194 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 240

Qy        254 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 300

Qy        314 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 373
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 GACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGT 360

Qy        374 CAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGC 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGC 420

Qy        434 CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTAC 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTAC 480

Qy        494 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 540

Qy        554 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 600

Qy        614 CCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCT 673
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        601 CCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTKCCGGAAAACAGCT 660

Qy        674 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 720

Qy        734 ACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 780

Qy        794 CCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGA 853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGA 840

Qy        854 AGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCG 913
              |||||||||||||||||||||||||||||:|||:|:|||||:|||||||||||||||||
Db        841 AGGCATTCCTCCTKACCAGCAGAGGTTGATYTTTKCYGGAAARCAGCTGGAAGATGGGCG 900

Qy        914 CACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAG 973
              |||||||||||||||:||||||||||||:|||||||||||||||:|| ||||||::|:||
Db        901 CACCCTGTCTGACTAMAACATCCAGAAARAGTCTACCCTGCACCKGGGGCTCCGKYTYAG 960

Qy        974 AGGTGGGATGCA-GATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC-CCTCGAAGTG 1031
              |:| |||:||:| |||:||:||::|:|:|||||:|:|  | | ||::|| | ||||||||
Db        961 ARGGGGGRTGSARGATYTTYGTKRARAMCCTGAMTKGGKAAAACAWYACTCATCGAAGTG 1020

Qy       1032 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              | ||||||||::| ||:|:|||| | :| :  || |::||| | || | | |
Db       1021 GGGCCGAGTGRMAACAWTKAGAAATTTWAARGSAAAAWYCAAAAAAAAGGAAG 1073
```

```
RESULT 13
BU103666
LOCUS       BU103666               1618 bp    mRNA    linear   EST 12-AUG-2003
DEFINITION  SCCCHR1002A02 Saccharum officinarum mRNA (Nogueira,F.T.S) Saccharum
            officinarum cDNA, mRNA sequence.
ACCESSION   BU103666
```

```
VERSION     BU103666.1  GI:32814997
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 1618)
  AUTHORS   Nogueira,F.T.S., de Rosa,V.E. Jr., Menossi,M., Ulian,E.C. and
            Arruda,P.
  TITLE     RNA expression profiles and data mining of sugarcane response to
            low temperature
  JOURNAL   Plant Physiol. 132 (4), 1811-1824 (2003)
   PUBMED   12913139
COMMENT     Contact: Nogueira FTS
            Bioinformatics Lab
            Organization for Nucleotide Sequencing and Analysis
            C.P. 6176, Campinas, SP 13083-970, Brazil
            Tel: 55 19 37881101
            Fax: 55 19 37881089
            Email: tebaldi@unicamp.br.
FEATURES             Location/Qualifiers
     source          1. .1618
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone_lib="Saccharum officinarum mRNA (Nogueira,F.T.S)"
ORIGIN

  Query Match             45.4%;  Score 995.4;  DB 5;  Length 1618;
  Best Local Similarity   82.5%;  Pred. No. 6e-261;
  Matches 1140;  Conservative    0;  Mismatches  241;  Indels    0;  Gaps    0;

Qy          69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
               ||||||||||||||| ||||| || ||||| ||||||||||||||| ||||||||| ||
Db           1 ATGCAGATCTTCGTTAAGACCCTCACTGGCAAGACCATCACCCTTGAGGTTGAGTCCTCA 60

Qy         129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
               |||||  || ||  |||||||||| ||||||| || |||||||||| |||||  || |||
Db          61 GACACTATTGACAATGTCAAGGCTAAGATCCAGGACAAGGAAGGCATTCCTCCGGATCAG 120

Qy         189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
               |||||||||||||||||| || ||||| || ||||| || |||||  ||||||||||
Db         121 CAGAGGCTGATCTTTGCTGGCAAGCAGCTCGAGGATGGCCGTACCCTAGCTGACTACAAC 180

Qy         249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
               |||||||  |||||||||| |||||||||||  | ||| |||| || |||||  ||||||
Db         181 ATCCAGAAGGAGTCCACCCTCCACCTGGTGCTCAGGCTCCGAGGAGGCATGCAGATCTTC 240

Qy         309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
               ||  |||||  |||||||||||||| |||||  |||||||| |||  |   ||||||  |||||
Db         241 GTCAAGACCCTCACTGGCAAGACTATCACGCTTGAGGTTGAGTCTTCTGACACGATCGAC 300

Qy         369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
               |||||  ||  ||  ||||||||||||||||||  ||  ||  ||  ||  |||||||||  |   | |||
Db         301 AACGTGAAGGCCAAGATCCAGGACAAGGAGGGAATCCCCCCAGACCAGCAGCGTCTCATC 360

Qy         429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
               ||  ||  ||  ||||||||  ||||||||  ||||||||  ||||||||||||||||||  |||
Db         361 TTCGCTGGCAAGCAGCTTGAAGATGGCCGCACCCTCGCTGACTACAACATCCAGAAGGAG 420

Qy         489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
               ||  ||  ||  |||||  ||||||  |  |||||  |||||  |||||  |||||  ||  ||||||||
Db         421 TCGACTCTTCACCTTGTGCTCAGGCTCAGGGGTGGCATGCAAATCTTTGTCAAGACCCTC 480

Qy         549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
               |||||  ||||||||||||  |  |||||||||  ||    ||||||||  ||  |||||  |||||
Db         481 ACTGGCAAGACCATCACCTTGGAGGTGGAGTCCTCGGACACCATTGACAATGTGAAGGCG 540

Qy         609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
               ||||||||  ||  |||||  ||||||||  ||  ||  ||||||  |   |  |||||||||  ||  ||
Db         541 AAGATCCAGGACAAGGAGGGCATTCCCCCGGACCAGCAGCGTCTCATCTTTGCTGGCAAG 600

Qy         669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
               |||||  ||  |||||  ||||||||  ||||||||||||||||||  |||||||||  |  |||
Db         601 CAGCTCGAGGATGGCCGTACCCTAGCTGACTACAACATCCAGAAGGAGTCCACCCTCCAC 660

Qy         729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
               |||||  |||  |  |||||  ||  ||  ||||||||||||||  |||||  ||||||||||||||
Db         661 CTGGTGCTCAGGCTCAGGGGAGGCATGCAAATCTTCGTCAAGACCCTCACTGGCAAGACT 720

Qy         789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
               |||||  ||||||||||||  |  ||||||  |||||  |||||  ||  ||  |||||||||  |||
Db         721 ATCACGCTTGAGGTCGAGTCTTCTGACACGATCGACAACGTGAAGGCCAAGATCCAGGAC 780

Qy         849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
               |||||  ||  ||  ||  ||  |||||||||  |  |  |||||  ||  ||  ||||||||  ||  |||
Db         781 AAGGAGGGAATCCCCCCGGACCAGCAGCGTCTCATCTTCGCTGGCAAGCAGCTCGAGGAT 840

Qy         909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
               ||  ||||||||  ||||||||||||||||||  |||||  ||  ||  |||||  ||||||  |
Db         841 GGCCGCACCCTCGCTGACTACAACATCCAGAAGGAGTCGACTCTCCACCTTGTGCTCAGG 900

Qy         969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028
               |||||  |||||  ||||||||||||||  ||||||||  ||  ||  ||||||||||||  |  ||
Db         901 CTCAGGGGTGGCATGCAGATCTTCGTCAAGACCCTCACCGGCAAGACCATCACCTTGGAG 960

Qy        1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
               ||||||  |     ||||||||||  |||||  |||||  ||||||||  ||||||||  ||||||
Db         961 GTGGAGTCCTCGGACACCATTGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCATC 1020

Qy        1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
               ||  ||  |||||||||  |  |  ||||||||  ||  ||  |||||  ||  |||||
Db        1021 CCCCCAGACCAGCAGCGTCTTATCTTTGCTGGCAAGCAGCTTGAGGATGGCCGCACCCTA 1080

Qy        1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
               |  ||||||||||||||||||  |||||||||  |  ||||||||||||||  ||  |  |||||
Db        1081 GCAGACTACAACATCCAGAAGGAGTCCACCCTTCACCTGGTGCTCCGCCTTCGCGGTGGT 1140

Qy        1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
               ||||||||||||||  ||||||||  ||  ||  ||||||||||  ||  |||||||||  |    |
Db        1141 ATGCAGATCTTCGTCAAGACCCTCACCGGCAAGACCATCACCCTGGAGGTGGAGTCCTCT 1200

Qy        1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
               ||||||||  ||  |||||  |||||  ||||||||  ||||||||  ||||||||  ||  ||  |||
Db        1201 GACACCATCGACAATGTGAAGGCGAAGATCCAGGACAAGGAGGGCATCCCCCCGGACCAG 1260

Qy        1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
               |||  |  |||||  ||  ||  ||||||||  |||||  |||||||||  |  |||||||||
Db        1261 CAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCACCCTGGCAGACTACAAC 1320

Qy        1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448
               ||||||||  ||||||||  ||  ||||||||||||||||||  |  |||||  |  |  ||  |
Db        1321 ATCCAGAAGGAGTCCACTCTCCACCTGGTGCTCCGTCTCCGTGGTGGCCAGTAAGTCCTG 1380

Qy        1449 G 1449
               |
Db        1381 G 1381


RESULT 14
AL533345
LOCUS       AL533345               1117 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL533345 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN003YB14 5-PRIME, mRNA sequence.
ACCESSION   AL533345
VERSION     AL533345.3  GI:45708245
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1117)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
```

```
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31071177.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN003DA07QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN003YB14"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             45.1%;  Score 987.6;  DB 1;  Length 1117;
  Best Local Similarity   95.6%;  Pred. No. 7.2e-259;
  Matches 1050;  Conservative    7;  Mismatches   36;  Indels    5;  Gaps    4;

Qy         14 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 73
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db         10 GCCGGGATTTGGGTCGCAGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 69

Qy         74 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 129

Qy        134 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 189

Qy        194 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 249

Qy        254 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 309

Qy        314 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 373
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        310 GACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGT 369

Qy        374 CAAAGCAAAGATCCA--GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 431
              |||||||||||||||  |    ||||||||||||||||||||||||||||||||||||
Db        370 CAAAGCAAAGATCCACGGACACAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 429

Qy        432 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 489

Qy        492 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 549

Qy        552 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 609

Qy        612 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 669

Qy        672 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 729

Qy        732 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 789

Qy        792 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAG 851
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAGGCAAAGATCCAAGACAAG 849

Qy        852 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 911
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        850 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAARCAGCTGGAAGATGGG 909

Qy        912 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 969

Qy        972 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTG 1031
              |:|||||||||||||||||||||||||||||||||||:| ||||||| |||||||||||
Db        970 ARAGGTGGGATGCAGATCTTCGTGAAGACCTTGACTGGTAARA-CATCACTCTCGAAGTG 1028

Qy       1032 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1091
              |:| |||||||||||| |:|||||:|||: |||      | |  | | |  |  |
Db       1029 GRG-CGAGTGACACCA-TKAGAATGTSAAGSAAAATCCAAGAMAGGAAGSATCCYCCTGA 1086

Qy       1092 CCTGACCAGCAGAGGTTG 1109
              || |   ||  ||  |||
Db       1087 CCAGCAAAGTTGAYYTTG 1104


RESULT 15
CL952190
LOCUS       CL952190                1854 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsJRUA000005 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL952190
VERSION     CL952190.1  GI:52364199
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1854)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .1854
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
```

/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Query Match 44.9%; Score 984.8; DB 10; Length 1854;
Best Local Similarity 82.1%; Pred. No. 5e-258;
Matches 1133; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

```
Qy        56 TCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGA 115
             |||    | | |  |||||||||| ||||||||  |||| || ||||| ||||||||||
Db        66 TCGCGTCCTCAAGATGCAGATCTTTGTGAAGACATTGACCGGCAAGACTATCACCCTCGA 125

Qy       116 GGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 175
             ||| ||| ||  |||||||||||| |||||||||| |||||||||||||||||| |||||
Db       126 GGTGGAGTCCTCTGACACCATCGATAATGTCAAGGCTAAGATCCAAGATAAGGAGGGCAT 185

Qy       176 CCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCT 235
             ||| || ||||||||| | || ||||| ||||| || |||||||| |||||   | |||||
Db       186 CCCCCCGGACCAGCAGCGTCTCATCTTCGCTGGCAAGCAGCTGGAGGATGGCAGGACCCT 245

Qy       236 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 295
               |||||||||||||||||| ||||| ||||| ||||| || ||||| ||| | |||||
Db       246 TGCTGACTACAACATCCAGAAGGAGTCGACCCTTCACCTTGTCCTCCGCCTCCGTGGTGG 305

Qy       296 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAG 355
             |||||  |||| || |||| || || |||||||| ||||||||||||||||||| |
Db       306 CATGCAGATCTTTGTCAAGACTCTGACCGGCAAGACTATCACCCTTGAGGTGGAGTCTTC 365

Qy       356 TGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCA 415
             |||||||||||| |||||||| || ||||||||||||||| || ||||| || || |||||
Db       366 TGACACCATCGACAACGTCAAGGCCAAGATCCAGGACAAAGAGGGCATCCCCCCAGACCA 425

Qy       416 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 475
             ||||  |  | ||||| ||||| |||||||||| |||||  | |||||  |||||||||
Db       426 GCAGCGTCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCAGGACCCTTGCTGACTACAA 485

Qy       476 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 535
             |||||||||  ||||| ||||| ||||| || ||||| ||| |  ||||| ||||||||||
Db       486 CATCCAGAAGGAGTCCACCCTCCACCTTGTCCTCCGCCTCCGTGGTGGCATGCAGATCTT 545

Qy       536 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 595
               || ||||| ||||| || |||||||||||||||||||||||  |   ||||||||||||
Db       546 TGTCAAGACACTGACCGGCAAGACCATCACCCTCGAGGTGGAATCTTCTGACACCATCGA 605

Qy       596 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGAT 655
               || |||||||| |||||||| || ||||| |||||||| || || |||||| |  | ||
Db       606 CAACGTCAAGGCCAAGATCCAGGACAAGGAGGGCATTCCCCCGGACCAGCAGCGTCTCAT 665

Qy       656 CTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGA 715
             |||||||||  || ||||| || || ||  | |||||  |||||||||||||||||||| ||
Db       666 CTTTGCCGGCAAGCAGCTTGAGGACGGCAGGACCCTTGCTGACTACAACATCCAGAAGGA 725

Qy       716 GTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 775
             |||  ||  | ||||| || |||||||||||  || || ||||||||||||||||||| ||
Db       726 GTCAACGCTTCACCTTGTCCTCCGTCTCAGGGGAGGCATGCAAATCTTCGTGAAGACTCT 785

Qy       776 CACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGC 835
               || |||||||||||||||| ||||| ||| |   ||| || ||||| || ||||| ||
Db       786 GACCGGCAAGACCATCACCCTCGAGGTGGAGTCTTCTGATACCATCGACAATGTCAAGGC 845

Qy       836 AAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 895
             ||||||||  |||||||| |||||||| || ||||||||  |  | ||||||| || ||
Db       846 CAAGATCCAGGACAAGGAGGGCATTCCCCCGGACCAGCAGCGCCTCATCTTTGCTGGCAA 905

Qy       896 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 955
             |||||||||  |||||  | |||||  |||||||||||||||||||| |||||  ||||| ||
Db       906 GCAGCTGGAGGATGGCAGGACCCTTGCTGACTACAACATCCAGAAGGAGTCCACCCTCCA 965

Qy       956 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1015
             |||  |||||||| ||  | |||||  |||||||||| || |||||||| ||. || |||||
Db       966 CCTTGTGCTCCGCCTTCGTGGTGGTATGCAGATCTTTGTCAAGACCCTCACAGGCAAGAC 1025

Qy      1016 CATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1075
             |||||||| || || |||     |||||||| || || |||||||| |||||||| ||
Db      1026 CATCACCCTGGAGGTTGAGAGCTCGGACACCATCGACAACGTCAAGGCCAAGATCCAGGA 1085

Qy      1076 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 1135
             ||||||  |||||||| || |||||||| |  | |||||  |||||  || |||||  || ||
Db      1086 CAAGGAGGGCATCCCCCCAGACCAGCAGCGTCTCATCTTCGCCGGCAAGCAGCTCGAGGA 1145

Qy      1136 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCG 1195
             |||  || |||||   | |||||||||||||||||  |||||  |||  |  |||||||||| ||
Db      1146 TGGCCGCACCCTCGCCGACTACAACATCCAGAAGGAGTCTACCCTCCACCTGGTGCTTCG 1205

Qy      1196 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1255
             ||||  |  |||||  |||||||||||||||||||| |||||||  |||||||||||| | ||
Db      1206 TCTCCGTGGTGGTATGCAGATCTTCGTGAAGACCTTGACTGGGAAGACCATCACTTTGGA 1265

Qy      1256 GGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1315
             |||  |||      |||||||||| ||||| ||||| |||||||| |||||||| || ||
Db      1266 GGTTGAGAGCTCCGACACCATTGATAATGTGAAGGCCAAGATCCAGGACAAGGAGGGGAT 1325

Qy      1316 CCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1375
               || || || |||||| |  ||||||| ||||| || |||||||| ||||||||||||||
Db      1326 TCCCCCAGACCAGCAGCGTCTGATCTTCGCTGGCAAGCAGCTGGAGGATGGACGCACCCT 1385

Qy      1376 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1435
               | |||||||||||||||| |||||||||| |||||||||||||  ||   | |||||
Db      1386 CGCCGACTACAACATCCAGAAGGAGTCCACCCTCCACCTGGTGCTCCGCCTCCGTGGTGG 1445
```

Search completed: March 7, 2006, 00:53:47
Job time : 8813 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 20:30:40 ; Search time 10737 Seconds
(without alignments)
11604.810 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 100

Total number of hits satisfying chosen parameters: 223

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2192	100.0	2192	6	CS097362	CS097362 Sequence
	2	2192	100.0	2192	6	E12458	E12458 cDNA encodi
	3	2192	100.0	2192	6	E12459	E12459 cDNA encodi
	4	2192	100.0	2192	6	E12605	E12605 cDNA encodi
	5	2192	100.0	2192	8	AB009010	AB009010 Homo sapi
	6	2191	99.9	2222	8	BC039193	BC039193 Homo sapi
	7	2177	99.3	2201	6	CQ979237	CQ979237 Sequence
	8	2177	99.3	2201	8	BC000449	BC000449 Homo sapi
	9	2177	99.3	2210	6	CQ979236	CQ979236 Sequence
	10	2165	98.8	2187	8	BC080583	BC080583 Homo sapi
	11	2069	94.4	2191	8	AK098827	AK098827 Homo sapi
	12	1940	88.5	2207	8	BC093445	BC093445 Homo sapi
	13	1906	87.0	2191	8	AK129749	AK129749 Homo sapi
c	14	1873	85.4	121984	8	AC126309	AC126309 Homo sapi
	15	1854	84.6	2058	8	AB089613	AB089613 Homo sapi
	16	1633	74.5	6605	8	HUMPOLYUBI	D63791 Human UbC g
	17	1284	58.6	2291	8	AB003730	AB003730 Homo sapi
	18	1283	58.5	2309	6	AX337205	AX337205 Sequence
	19	1283	58.5	2309	8	HUMUBI13	M26880 Human ubiqu
	20	1154	52.6	1864	8	AK097415	AK097415 Homo sapi
	21	1123	51.2	1728	8	BC008955	BC008955 Homo sapi
	22	1103	50.3	1122	8	BC014880	BC014880 Homo sapi
	23	1017	46.4	2784	8	AK092983	AK092983 Homo sapi
	24	837	38.2	3996	8	AB209782	AB209782 Homo sapi
	25	827	37.7	1168	8	AF052164	AF052164 Homo sapi
	26	824	37.6	824	8	HUMUB	M17597 Human poly-
	27	735	33.5	796	8	HSM801035	AL117514 Homo sapi
	28	719	32.8	1300	6	AX714536	AX714536 Sequence
	29	719	32.8	1300	8	AK056939	AK056939 Homo sapi
	30	716	32.7	1201	10	BV177540	BV177540 sqnm95284
c	31	682	31.1	1251	10	BV176676	BV176676 sqnm90309
c	32	682	31.1	1251	10	BV176677	BV176677 sqnm90313
c	33	682	31.1	1251	10	BV176678	BV176678 sqnm90314
c	34	682	31.1	1251	10	BV177180	BV177180 sqnm93250
c	35	682	31.1	1251	10	BV177599	BV177599 sqnm95667
c	36	682	31.1	1251	10	BV180128	BV180128 sqnm10811
c	37	682	31.1	1251	10	BV180129	BV180129 sqnm10812
	38	632	28.8	718	8	AY819759	AY819759 Homo sapi
	39	571	26.0	1082	8	AK026846	AK026846 Homo sapi
	40	556	25.4	2286	8	AB089614	AB089614 Pan trogl
	41	554	25.3	1110	10	BV177342	BV177342 sqnm94123
	42	554	25.3	1110	10	BV179342	BV179342 sqnm10493
c	43	456	20.8	1063	10	BV176674	BV176674 sqnm90307
c	44	456	20.8	1063	10	BV176675	BV176675 sqnm90308
	45	449	20.5	547	6	BD237119	BD237119 Compounds
	46	449	20.5	547	6	AR225519	AR225519 Sequence
	47	449	20.5	547	6	AR562942	AR562942 Sequence
	48	449	20.5	547	6	AX321640	AX321640 Sequence
	49	433	19.8	460	6	BD265164	BD265164 Compounds
	50	433	19.8	460	6	AR401150	AR401150 Sequence
	51	433	19.8	460	6	AX192596	AX192596 Sequence
	52	430	19.6	494	6	AR225588	AR225588 Sequence
	53	430	19.6	494	6	AR563011	AR563011 Sequence
	54	430	19.6	494	6	AX321744	AX321744 Sequence
	55	426	19.4	483	6	BD237089	BD237089 Compounds
	56	426	19.4	483	6	AR225489	AR225489 Sequence
	57	426	19.4	483	6	AR562912	AR562912 Sequence
	58	426	19.4	483	6	AX321610	AX321610 Sequence
	59	411	18.8	412	6	AR563410	AR563410 Sequence
	60	410	18.7	434	6	AR225561	AR225561 Sequence
	61	410	18.7	434	6	AR562984	AR562984 Sequence
	62	410	18.7	434	6	AX321717	AX321717 Sequence
	63	402	18.3	470	6	CQ696720	CQ696720 Sequence
	64	370	16.9	518	6	BD264767	BD264767 Compositi
	65	370	16.9	518	6	AR238170	AR238170 Sequence
	66	370	16.9	518	6	AR257711	AR257711 Sequence
	67	370	16.9	518	6	AR283757	AR283757 Sequence
	68	370	16.9	518	6	AR443431	AR443431 Sequence
	69	370	16.9	518	6	AR478509	AR478509 Sequence
	70	370	16.9	518	6	AR640757	AR640757 Sequence
	71	370	16.9	518	6	AX366385	AX366385 Sequence
	72	327	14.9	400	6	CQ667521	CQ667521 Sequence
	73	327	14.9	401	6	CQ688969	CQ688969 Sequence
	74	323	14.7	472	6	CQ697121	CQ697121 Sequence
	75	304	13.9	389	6	CQ690613	CQ690613 Sequence
	76	303	13.8	308	6	CQ684318	CQ684318 Sequence
	77	303	13.8	412	6	AX193093	AX193093 Sequence
	78	299	13.6	312	6	CQ689249	CQ689249 Sequence
	79	299	13.6	400	6	CQ676939	CQ676939 Sequence
	80	281	12.8	332	6	CQ704345	CQ704345 Sequence
	81	276	12.6	361	6	CQ686953	CQ686953 Sequence
	82	271	12.4	396	6	CQ688744	CQ688744 Sequence
	83	270	12.3	284	6	AR203067	AR203067 Sequence
	84	270	12.3	288	6	CQ667335	CQ667335 Sequence
	85	266	12.1	490	8	AY732488	AY732488 Homo sapi
	86	259	11.8	345	6	CQ698215	CQ698215 Sequence
	87	258	11.8	339	6	CQ686946	CQ686946 Sequence
	88	258	11.8	345	6	CQ705082	CQ705082 Sequence
	89	255	11.6	337	6	CQ685697	CQ685697 Sequence
	90	255	11.6	543	6	CQ923979	CQ923979 Sequence
	91	254	11.6	565	6	CQ924043	CQ924043 Sequence

92	248	11.3	585	6	BD264660	BD264660 Compositi
93	248	11.3	585	6	AR238063	AR238063 Sequence
94	248	11.3	585	6	AR257604	AR257604 Sequence
95	248	11.3	585	6	AR283650	AR283650 Sequence
96	248	11.3	585	6	AR443324	AR443324 Sequence
97	248	11.3	585	6	AR478402	AR478402 Sequence
98	248	11.3	585	6	AR640650	AR640650 Sequence
99	248	11.3	585	6	AX366278	AX366278 Sequence
c 100	244	11.1	467	6	CQ514028	CQ514028 Sequence
101	242	11.0	286	6	AX187556	AX187556 Sequence
102	242	11.0	329	6	CQ665885	CQ665885 Sequence
103	242	11.0	487	6	AX186495	AX186495 Sequence
104	238	10.9	337	6	CQ706133	CQ706133 Sequence
105	236	10.8	474	6	CQ683715	CQ683715 Sequence
106	235	10.7	286	6	CQ688311	CQ688311 Sequence
107	233	10.6	237	6	CQ682439	CQ682439 Sequence
108	232	10.6	232	6	CQ689616	CQ689616 Sequence
109	223	10.2	232	6	CQ704983	CQ704983 Sequence
110	218	9.9	403	6	CQ697815	CQ697815 Sequence
111	217	9.9	372	6	CQ680957	CQ680957 Sequence
112	213	9.7	249	6	CQ682331	CQ682331 Sequence
113	211	9.6	440	6	AR240521	AR240521 Sequence
114	210	9.6	261	6	CQ686061	CQ686061 Sequence
115	210	9.6	330	6	CQ663336	CQ663336 Sequence
116	207	9.4	228	6	AR606553	AR606553 Sequence
c 117	207	9.4	609	6	BD229494	BD229494 Human gen
118	204	9.3	384	6	CQ658851	CQ658851 Sequence
119	203	9.3	480	6	AX816038	AX816038 Sequence
120	203	9.3	504	6	AX816040	AX816040 Sequence
121	202	9.2	261	6	CQ691540	CQ691540 Sequence
122	202	9.2	490	6	CQ695149	CQ695149 Sequence
123	201	9.2	549	6	CQ923956	CQ923956 Sequence
124	200	9.1	216	6	CQ690997	CQ690997 Sequence
125	200	9.1	367	6	CQ668894	CQ668894 Sequence
126	192	8.8	294	6	CQ705824	CQ705824 Sequence
c 127	191	8.7	605	6	BD264710	BD264710 Compositi
c 128	191	8.7	605	6	AR238113	AR238113 Sequence
c 129	191	8.7	605	6	AR257654	AR257654 Sequence
c 130	191	8.7	605	6	AR283700	AR283700 Sequence
c 131	191	8.7	605	6	AR443374	AR443374 Sequence
c 132	191	8.7	605	6	AR478452	AR478452 Sequence
c 133	191	8.7	605	6	AR640700	AR640700 Sequence
c 134	191	8.7	605	6	AX366328	AX366328 Sequence
135	191	8.7	697	6	BD264715	BD264715 Compositi
136	191	8.7	697	6	AR238118	AR238118 Sequence
137	191	8.7	697	6	AR257659	AR257659 Sequence
138	191	8.7	697	6	AR283705	AR283705 Sequence
139	191	8.7	697	6	AR443379	AR443379 Sequence
140	191	8.7	697	6	AR478457	AR478457 Sequence
141	191	8.7	697	6	AR640705	AR640705 Sequence
142	191	8.7	697	6	AX366333	AX366333 Sequence
143	188	8.6	6340	6	AX207733	AX207733 Sequence
144	188	8.6	6411	6	AX207725	AX207725 Sequence
145	188	8.6	6411	6	AX207729	AX207729 Sequence
146	188	8.6	6436	6	AX207740	AX207740 Sequence
147	188	8.6	6459	6	AX207749	AX207749 Sequence
148	188	8.6	6459	6	AX207755	AX207755 Sequence
149	188	8.6	6607	6	AX207734	AX207734 Sequence
150	188	8.6	6678	6	AX207726	AX207726 Sequence
151	188	8.6	6703	6	AX207741	AX207741 Sequence
152	188	8.6	6726	6	AX207750	AX207750 Sequence
153	188	8.6	6726	6	AX207756	AX207756 Sequence
154	188	8.6	6850	6	AX207735	AX207735 Sequence
155	188	8.6	6921	6	AX207727	AX207727 Sequence
156	188	8.6	6946	6	AX207742	AX207742 Sequence
157	188	8.6	6969	6	AX207751	AX207751 Sequence
158	188	8.6	6969	6	AX207757	AX207757 Sequence
159	188	8.6	7053	6	AX207767	AX207767 Sequence
160	188	8.6	7053	6	AX207772	AX207772 Sequence
161	188	8.6	7053	6	AX207773	AX207773 Sequence
162	188	8.6	7093	6	AX207736	AX207736 Sequence
163	188	8.6	7164	6	AX207728	AX207728 Sequence
164	188	8.6	7189	6	AX207743	AX207743 Sequence
165	188	8.6	7212	6	AX207752	AX207752 Sequence
166	188	8.6	7212	6	AX207758	AX207758 Sequence
167	188	8.6	7248	6	AX207746	AX207746 Sequence
168	188	8.6	7512	6	AX207774	AX207774 Sequence
169	188	8.6	7563	6	AX207766	AX207766 Sequence
170	188	8.6	7563	6	AX207770	AX207770 Sequence
171	188	8.6	7563	6	AX207771	AX207771 Sequence
172	188	8.6	8022	6	AX207762	AX207762 Sequence
173	183	8.3	355	6	CQ679610	CQ679610 Sequence
174	182	8.3	321	6	CQ660774	CQ660774 Sequence
175	176	8.0	227	6	AX209208	AX209208 Sequence
176	176	8.0	335	6	AX198675	AX198675 Sequence
c 177	171	7.8	294	6	CQ484202	CQ484202 Sequence
178	167	7.6	228	6	AX207704	AX207704 Sequence
179	167	7.6	228	6	AX207719	AX207719 Sequence
180	161	7.3	273	6	CQ691600	CQ691600 Sequence
181	160	7.3	211	6	CQ705545	CQ705545 Sequence
182	150	6.8	201	10	BV207846	BV207846 sqnm22469
183	143	6.5	400	6	CQ663822	CQ663822 Sequence
184	143	6.5	1830	8	AB089615	AB089615 Gorilla g
185	142	6.5	201	10	BV200498	BV200498 sqnm20427
186	142	6.5	201	10	BV200925	BV200925 sqnm20542
c 187	137	6.2	240	6	AX198840	AX198840 Sequence
c 188	137	6.2	240	6	AX209367	AX209367 Sequence
189	137	6.2	2286	8	AB089616	AB089616 Pongo pyg
190	135	6.2	208	6	CQ700105	CQ700105 Sequence
191	133	6.1	133	6	CS025483	CS025483 Sequence
192	132	6.0	374	6	CQ689709	CQ689709 Sequence
193	126	5.7	251	10	BV168033	BV168033 sqnm7091
194	126	5.7	369	6	BD225145	BD225145 Medicinal
195	126	5.7	579	6	BD225146	BD225146 Medicinal
196	126	5.7	651	6	BD225149	BD225149 Medicinal
197	126	5.7	891	6	BD225147	BD225147 Medicinal
198	126	5.7	1020	6	BD225148	BD225148 Medicinal
199	126	5.7	1737	6	BD225150	BD225150 Medicinal
200	125	5.7	168	6	CQ659465	CQ659465 Sequence
201	125	5.7	255	6	AX184690	AX184690 Sequence
202	124	5.7	201	10	BV207847	BV207847 sqnm22469
c 203	123	5.6	636	6	BD264714	BD264714 Compositi
c 204	123	5.6	636	6	AR238117	AR238117 Sequence
c 205	123	5.6	636	6	AR257658	AR257658 Sequence
c 206	123	5.6	636	6	AR283704	AR283704 Sequence
c 207	123	5.6	636	6	AR443378	AR443378 Sequence
c 208	123	5.6	636	6	AR478456	AR478456 Sequence
c 209	123	5.6	636	6	AR640704	AR640704 Sequence
c 210	123	5.6	636	6	AX366332	AX366332 Sequence
c 211	121	5.5	184	6	AX322311	AX322311 Sequence
212	120	5.5	341	6	CQ689661	CQ689661 Sequence
213	112	5.1	112	6	CQ683199	CQ683199 Sequence
214	109	5.0	167	6	CQ660012	CQ660012 Sequence
c 215	104	4.7	795	6	A87049	A87049 Sequence 14
c 216	104	4.7	795	6	AR048461	AR048461 Sequence
c 217	104	4.7	795	6	AR309783	AR309783 Sequence
218	100	4.6	107	6	CQ668311	CQ668311 Sequence
219	100	4.6	201	10	BV207050	BV207050 sqnm22260
220	100	4.6	201	10	BV207845	BV207845 sqnm22469
221	100	4.6	201	10	BV207848	BV207848 sqnm22470
222	100	4.6	201	10	BV208263	BV208263 sqnm22597
223	100	4.6	201	10	BV208264	BV208264 sqnm22597

ALIGNMENTS

RESULT 1
CS097362
LOCUS CS097362 2192 bp DNA linear PAT 03-JUN-2005
DEFINITION Sequence 5 from Patent WO2005045070.
ACCESSION CS097362
VERSION CS097362.1 GI:66953811
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Lacroix,B., Krause,A., Puisieux,A. and Bachelot,T.
TITLE Method for prognosticating a breast cancer
JOURNAL Patent: WO 2005045070-A 5 19-MAY-2005;
BIOMERIEUX (FR); Centre Leon Berard (FR)

FEATURES Location/Qualifiers
source 1. .2192
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
Db    1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy   61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Db   61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy  121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Db  121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy  181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Db  181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy  241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Db  241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy  301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db  301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy  361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db  361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy  421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db  421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy  481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db  481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy  541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db  541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy  601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db  601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy  661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db  661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy  721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db  721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy  781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db  781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy  841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db  841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy  901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db  901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy  961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db  961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy 1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db 1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy 1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db 1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy 1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db 1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy 1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db 1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy 1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db 1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy 1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db 1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy 1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db 1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy 1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db 1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy 1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db 1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy 1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db 1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy 1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db 1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy 1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db 1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy 1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db 1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy 1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db 1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy 1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db 1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
```

```
Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 2
E12458
LOCUS E12458 2192 bp DNA linear PAT 27-APR-1998
DEFINITION cDNA encoding human polyubiquitin.
ACCESSION E12458
VERSION E12458.1 GI:3251291
KEYWORDS JP 1997000262-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2192)
AUTHORS Kato,M., Yamaguchi,T., Kin,N. and Sekine,S.
TITLE HUMAN UBIQUITIN-BOUND ENZYME E2 AND CDNA CODING THE SAME
JOURNAL Patent: JP 1997000262-A 2 07-JAN-1997;
SAGAMI CHEM RES CENTER
COMMENT OS Homo sapiens
PN JP 1997000262-A/2
PD 07-JAN-1997
PF 21-JUN-1995 JP 1995154615
PI KATO MASASHI, YAMAGUCHI TOMOKO, KIN NANJIYUN, SEKINE SHINGO PC
C12N15/09,C07H21/04,C07K14/46,C12N9/64,C12P21/02,C12Q1/68, PC
(C12N15/09,
PC C12R1:91),(C12N9/64,C12R1:19),(C12P21/02,C12R1:19); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH
FT source 1. .2192
FT /organism='Homo sapiens'
FT /cell_type='lymphoma'
FT /cell_line='U937'
FT /clone='HP00184'
FT CDS 69. .2126
FT /product='human polyubiquitin'.
FEATURES Location/Qualifiers
source 1. .2192
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
RESULT 3
E12459
LOCUS       E12459                  2192 bp    DNA     linear   PAT 27-APR-1998
DEFINITION  cDNA encoding human polyubiquitin.
ACCESSION   E12459
VERSION     E12459.1  GI:3251292
KEYWORDS    JP 1997000263-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2192)
  AUTHORS   Kato,M., Saeki,M., Kato,T., Kin,N. and Sekine,S.
  TITLE     POLYUBIQUITIN CDNA AND PRODUCTION OF LABELED UBIQUITIN
  JOURNAL   Patent: JP 1997000263-A 1 07-JAN-1997;
            SAGAMI CHEM RES CENTER
COMMENT     OS   Homo sapiens
            PN   JP 1997000263-A/1
            PD   07-JAN-1997
            PF   21-JUN-1995 JP 1995154616
            PI   KATO MASASHI, SAEKI MIHORO, KATO TAKAE, KIN NANJIYUN, PI
            SEKINE SHINGO
            PC   C12N15/09,C07H21/04,C07K19/00,C12P21/02,C12Q1/68,(C12N15/09,
            PC   C12R1:91),
            PC   (C12P21/02,C12R1:91);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2192
            FT                   /organism='Homo sapiens'
            FT                   /cell_type='lymphoma'
            FT                   /cell_line='U937'
            FT                   /clone='HP00184'
            FT   5'UTR           1. .68
            FT   CDS             69. .2126
            FT                   /product='human polyubiquitin' FT   3'UTR
            2127. .2192.
FEATURES            Location/Qualifiers
     source         1. .2192
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match          100.0%;  Score 2192;  DB 6;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
```

```
Qy       301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db       301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Db      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
RESULT 4
E12605
LOCUS       E12605                2192 bp    DNA     linear   PAT 27-APR-1998
DEFINITION  cDNA encoding human polyubiquitin.
ACCESSION   E12605
VERSION     E12605.1  GI:3251437
KEYWORDS    JP 1997037779-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2192)
  AUTHORS   Kato,M., Saeki,M. and Sekine,S.
  TITLE     UBIQUITIN BOND ENZYME E2 AND HUMAN CDNA CODING FOR THE SAME
  JOURNAL   Patent: JP 1997037779-A 2 10-FEB-1997;
```

```
            SAGAMI CHEM RES CENTER
COMMENT     OS   Homo sapiens (human)
            PN   JP 1997037779-A/2
            PD   10-FEB-1997
            PF   31-JUL-1995 JP 1995194376
            PI   KATO MASASHI, SAEKI MIHORO, SEKINE SHINGO
            PC   C12N9/00,C07H21/04,C07K14/46,C07K19/00,C12N15/09//C12N1/21, PC
              (C12N9/00,
            PC   C12R1:19);
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .2192
            FT                   /organism='Homo sapiens'
            FT                   /tissue_type='Lymphoma'
            FT                   /cell_line='U937'
            FT                   /clone='HP00184'
            FT   5'UTR           1. .68
            FT   CDS             69. .2126
            FT                   /product='Human polyubiquitin' FT   3'UTR
                 2127. .2192.
FEATURES             Location/Qualifiers
     source          1. .2192
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2192;  DB 6;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
```

```
Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
RESULT 5
AB009010
LOCUS       AB009010               2192 bp    mRNA    linear   PRI 26-JUL-2001
DEFINITION  Homo sapiens mRNA for polyubiquitin UbC, complete cds.
ACCESSION   AB009010
VERSION     AB009010.1  GI:2647407
KEYWORDS    polyubiquitin UbC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,N., Yamaguchi,T., Sekine,S., Saeki,M., Iwamuro,S. and Kato,S.
  TITLE     Cloning of human polyubiquitin cDNAs and a ubiquitin-binding assay
            involving its in vitro translation product
  JOURNAL   J. Biochem. 125, 35-39 (1998)
REFERENCE   2  (bases 1 to 2192)
  AUTHORS   Kato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-1997) Seishi Kato, Research Institute of National
            Rehabilitation Center for the Disabled, Department of
            Rehabilitation Engineering; 4-1 Namiki, Tokorozawa, Saitama
            359-8555, Japan (E-mail:seishi@rehab.go.jp,
            Tel:042-995-3100(ex.2568), Fax:042-995-3132)
FEATURES             Location/Qualifiers
     source          1. .2192
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HP00184"
     gene            join(31. .48,347,1022,1325,1898)
                     /gene="UbC2"
     variation       31
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="a"
     variation       48
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="t"
     gene            69. .2126
                     /gene="UbC1"
     CDS             69. .2126
                     /gene="UbC1"
                     /codon_start=1
                     /product="polyubiquitin UbC"
                     /protein_id="BAA23632.1"
                     /db_xref="GI:2647408"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGV"
     variation       347
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="c"
     variation       1022
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="t"
     variation       1325
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="c"
     variation       1898
                     /gene="UbC2"
                     /note="polymorphic variant"
                     /replace="a"
     polyA_signal    2173. .2178
ORIGIN
```

```
  Query Match           100.0%;  Score 2192;  DB 8;  Length 2192;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
```

```
Qy       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db       841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db       901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db       961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db      1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db      1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db      1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db      1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db      1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db      1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db      1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db      1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db      1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db      1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db      1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db      1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db      1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db      1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db      1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db      1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db      1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db      2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Db      2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db      2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 6
BC039193

```
LOCUS       BC039193               2222 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens ubiquitin C, mRNA (cDNA clone MGC:14624
            IMAGE:4076286), complete cds.
ACCESSION   BC039193
VERSION     BC039193.1  GI:24657521
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2222)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
```

```
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2222)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 21 Row: g Column: 17.
FEATURES             Location/Qualifiers
     source          1. .2222
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:14624 IMAGE:4076286"
                     /tissue_type="Brain, primitive neuroectodermal"
                     /clone_lib="NIH_MGC_56"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1. .2222
                     /gene="UBC"
                     /note="synonym: HMG20"
                     /db_xref="GeneID:7316"
                     /db_xref="MIM:191340"
     CDS             69. .2126
                     /gene="UBC"
                     /codon_start=1
                     /product="ubiquitin C"
                     /protein_id="AAH39193.1"
                     /db_xref="GI:24657522"
                     /db_xref="GeneID:7316"
                     /db_xref="MIM:191340"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGV"
ORIGIN

  Query Match             99.9%;  Score 2191;  DB 8;  Length 2222;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61

Qy         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121

Qy        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181

Qy        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241

Qy        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301

Qy        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361

Qy        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421

Qy        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481

Qy        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541

Qy        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601

Qy        602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661

Qy        662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 721

Qy        722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 CTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781

Qy        782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 CAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGAT 841

Qy        842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 CCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 901

Qy        902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 961

Qy        962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021
```

```
Db        962 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1021

Qy       1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081

Db       1022 CCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1081

Qy       1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141

Db       1082 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 1141

Qy       1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201

Db       1142 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAG 1201

Qy       1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261

Db       1202 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGA 1261

Qy       1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321

Db       1262 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCC 1321

Qy       1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381

Db       1322 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 1381

Qy       1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441

Db       1382 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCA 1441

Qy       1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501

Db       1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501

Qy       1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561

Db       1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561

Qy       1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621

Db       1562 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1621

Qy       1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681

Db       1622 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAA 1681

Qy       1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741

Db       1682 GACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGT 1741

Qy       1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801

Db       1742 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1801

Qy       1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861

Db       1802 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1861

Qy       1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921

Db       1862 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1921

Qy       1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981

Db       1922 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 1981

Qy       1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041

Db       1982 CCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 2041

Qy       2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101

Db       2042 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGT 2101

Qy       2102 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161

Db       2102 CCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGC 2161

Qy       2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192

Db       2162 ACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 7
CQ979237

```
LOCUS       CQ979237               2201 bp    DNA     linear   PAT 19-JAN-2005
DEFINITION  Sequence 3 from Patent WO2005001135.
ACCESSION   CQ979237
VERSION     CQ979237.1  GI:57976490
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elander,M.G., Ma,X.J. and Baer,T.M.
  TITLE     3' biased detection of nucleic acids
  JOURNAL   Patent: WO 2005001135-A 3 06-JAN-2005;
            Arcturus Bioscience, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2201
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match 99.3%; Score 2177; DB 6; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75

Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135

Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195

Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255

Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315

Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375

Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435

Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495

Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
```

```
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540
Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600
Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660
Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720
Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780
Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840
Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900
Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960
Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020
Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080
Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140
Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200
Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260
Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320
Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380
Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440
Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500
Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560
Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620
Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680
Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740
Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800
Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860
Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920
Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980
Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040
Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100
Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160
Qy       2176 AAAGTTGTTGCATTCCC 2192
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 8
BC000449
LOCUS BC000449 2201 bp mRNA linear PRI 06-NOV-2003
DEFINITION Homo sapiens ubiquitin C, mRNA (cDNA clone IMAGE:2821375), partial cds.
ACCESSION BC000449
VERSION BC000449.2 GI:38197156
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2201)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,

```
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2201)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-2000) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Nov 6, 2003 this sequence version replaced gi:12653358.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 1 Row: f Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 20149305.
FEATURES            Location/Qualifiers
     source         1. .2201
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2821375"
                    /tissue_type="Lung, small cell carcinoma"
                    /clone_lib="NIH_MGC_7"
                    /lab_host="DH10B-R"
                    /note="Vector: pOTB7"
     gene           <1. .2201
                    /gene="UBC"
                    /note="synonym: HMG20"
                    /db_xref="GeneID:7316"
                    /db_xref="MIM:191340"
     CDS            <1. .2111
                    /gene="UBC"
                    /codon_start=3
                    /product="UBC protein"
                    /protein_id="AAH00449.2"
                    /db_xref="GI:38197157"
                    /db_xref="GeneID:7316"
                    /db_xref="MIM:191340"
                    /translation="GFGSRFLFVDRCDRHLTMQIFVKTLTGKTITLEVEPSDTIENVK
                    AKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLT
                    GKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTL
                    HLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ
                    LEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQ
                    DKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTI
                    TLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVL
                    RLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDG
                    RTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG
                    IPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEV
                    EPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
                    GMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLS
                    DYNIQKESTLHLVLRLRGGV"
     misc_feature   54. .275
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   282. .503
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   510. .731
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   738. .959
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   966. .1187
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   1194. .1415
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   1422. .1643
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   1650. .1871
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
     misc_feature   1878. .2099
                    /gene="UBC"
                    /note="ubiquitin; Region: Ubiquitin family. This family
                    contains a number of ubiquitin-like proteins: SUMO (smt3
                    homologue), Nedd8, Elongin B, Rub1"
                    /db_xref="CDD:pfam00240"
ORIGIN

  Query Match           99.3%;  Score 2177;  DB 8;  Length 2201;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180
```

```
Qy      196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy      256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy      316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy      376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy      436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy      496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy      556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy      616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy      676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy      736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy      796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy      856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy      916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy      976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy     1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy     1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy     1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy     1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy     1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy     1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy     1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy     1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy     1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy     1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy     1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy     1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy     1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy     1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy     1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy     1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy     1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy     2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy     2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy     2176 AAAGTTGTTGCATTCCC 2192
            |||||||||||||||||
Db     2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 9
CQ979236
LOCUS CQ979236 2210 bp DNA linear PAT 19-JAN-2005
DEFINITION Sequence 2 from Patent WO2005001135.
ACCESSION CQ979236
VERSION CQ979236.1 GI:57976489
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elander,M.G., Ma,X.J. and Baer,T.M.
  TITLE     3' biased detection of nucleic acids
  JOURNAL   Patent: WO 2005001135-A 2 06-JAN-2005;
            Arcturus Bioscience, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2210
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.3%;  Score 2177;  DB 6;  Length 2210;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 69

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 129

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 189

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 249

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 309

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 369

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 429

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 489

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 549

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 609

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 669

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 729

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 789

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        790 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 849

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 909

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 969

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1029

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1089

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1149

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1209

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1269

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1329

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1389

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1449

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1509

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1569

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1629

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1689

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1749

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1809

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1810 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1869

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1870 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1929
```

```
Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1930 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1989

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1990 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2049

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2109

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2110 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2169

Qy       2176 AAAGTTGTTGCATTCCC 2192
             |||||||||||||||||
Db       2170 AAAGTTGTTGCATTCCC 2186


RESULT 10
BC080583
LOCUS       BC080583               2187 bp    mRNA    linear   PRI 07-SEP-2004
DEFINITION  Homo sapiens cDNA clone IMAGE:2822684, partial cds.
ACCESSION   BC080583
VERSION     BC080583.1  GI:51873889
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2187)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2187)
  AUTHORS   Director MGC Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-2004) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 59 Row: a Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 34304116.
FEATURES             Location/Qualifiers
     source          1. .2187
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2822684"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             <1. .2099
                     /codon_start=3
                     /product="Unknown (protein for IMAGE:2822684)"
                     /protein_id="AAH80583.1"
                     /db_xref="GI:51873890"
                     /translation="RFLFVDRCDRHLTMQIFVKTLTGKTITLEVEPSDTIENVKAKIQ
                     DKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTI
                     TLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVL
                     RLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDG
                     RTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG
                     IPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEV
                     EPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
                     GMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLS
                     DYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPD
                     QQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSD
                     TIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQI
                     FVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNI
                     QKESTLHLVLRLRGGV"
ORIGIN

  Query Match           98.8%;  Score 2165;  DB 8;  Length 2187;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360
```

```
Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 720

Qy        748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 807
Db        721 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 780

Qy        808 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 867
Db        781 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 840

Qy        868 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 927
Db        841 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 900

Qy        928 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 987
Db        901 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 960

Qy        988 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 1047
Db        961 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 1020

Qy       1048 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1107
Db       1021 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1080

Qy       1108 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 1167
Db       1081 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 1140

Qy       1168 AAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 1227
Db       1141 AAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 1200

Qy       1228 CCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCA 1287
Db       1201 CCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCA 1260

Qy       1288 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTG 1347
Db       1261 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTG 1320

Qy       1348 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1407
Db       1321 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1380

Qy       1408 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1467
Db       1381 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1440

Qy       1468 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1527
Db       1441 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1500

Qy       1528 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1587
Db       1501 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1560

Qy       1588 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1647
Db       1561 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1620

Qy       1648 TCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1707
Db       1621 TCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1680

Qy       1708 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1767
Db       1681 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1740

Qy       1768 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1827
Db       1741 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1800

Qy       1828 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 1887
Db       1801 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 1860

Qy       1888 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1947
Db       1861 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1920

Qy       1948 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 2007
Db       1921 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 1980

Qy       2008 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2067
Db       1981 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2040

Qy       2068 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2127
Db       2041 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2100

Qy       2128 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2187
Db       2101 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2160

Qy       2188 TTCCC 2192
Db       2161 TTCCC 2165
```

RESULT 11
AK098827
LOCUS AK098827 2191 bp mRNA linear PRI 12-SEP-2003
DEFINITION Homo sapiens cDNA FLJ25961 fis, clone TST09043, highly similar to Polyubiquitin UbC.
ACCESSION AK098827
VERSION AK098827.1 GI:21758959
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Suzuki,Y., Hata,H., Nakagawa,K.,

```
            Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B.,
            Nagai,K., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2191)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
            full insert sequencing: RAB and Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .2191
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TST09043"
                     /tissue_type="testis"
                     /clone_lib="TST"
                     /note="cloning vector: pME18SFL3"
ORIGIN

  Query Match             94.4%;  Score 2069;  DB 8;  Length 2191;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2169;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 80

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 140

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 200

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 260

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 320

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 380

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 440

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 500

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 560

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        561 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 620

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 680

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 740

Qy        742 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 800

Qy        802 TCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTC 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 TCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTC 860

Qy        862 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGT 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGT 920

Qy        922 CTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 CTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 980

Qy        982 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTG 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTG 1040

Qy       1042 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 1100

Qy       1102 AGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACA 1161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 AGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACA 1160

Qy       1162 TCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 1221
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 1220

Qy       1222 TGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGA 1281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1221 TGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGA 1280

Qy       1282 ATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCT 1341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1281 ATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCT 1340

Qy       1342 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 1401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1341 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 1400

Qy       1402 CCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 1461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1401 CCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 1460

Qy       1462 CTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAA 1521
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1461 CTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAA 1520

Qy       1522 AGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAAC 1581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1521 AGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAAC 1580

Qy       1582 AGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACC 1641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1581 AGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACC 1640

Qy       1642 TGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 1701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1641 TGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 1700
```

```
Qy       1702 TCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 1761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1701 TCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 1760

Qy       1762 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 1821
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1761 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 1820

Qy       1822 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 1881
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1821 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 1880

Qy       1882 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGG 1941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1881 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGG 1940

Qy       1942 TGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCC 2001
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1941 TGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCC 2000

Qy       2002 CTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 2061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 CTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 2060

Qy       2062 CTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 2121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 CTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 2120

Qy       2122 TCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTT 2181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 TCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTT 2180

Qy       2182 GTTGCATTCCC 2192
             ||||||||||||
Db       2181 GTTGCATTCCC 2191
```

RESULT 12
BC093445

```
LOCUS       BC093445               2207 bp    mRNA    linear   PRI 13-APR-2005
DEFINITION  Homo sapiens cDNA clone IMAGE:5284302, partial cds.
ACCESSION   BC093445
VERSION     BC093445.1  GI:62204947
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2207)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  CONSRTM   Mammalian Gene Collection Program Team
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2207)
  AUTHORS   .
  CONSRTM   NIH MGC Project
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2005) National Institutes of Health, Mammalian
            Gene Collection (MGC), Bethesda, MD 20892-2590, USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 168 Row: a Column: 20
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 34304116.
FEATURES             Location/Qualifiers
     source          1. .2207
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5284302"
                     /tissue_type="Brain, hippocampus"
                     /clone_lib="NIH_MGC_95"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript"
     CDS             <1. .2125
                     /codon_start=2
                     /product="Unknown (protein for IMAGE:5284302)"
                     /protein_id="AAH93445.1"
                     /db_xref="GI:62204948"
                     /translation="VPSQPGFGSQFLFVDCCDRHLTMQIFVKTLTGKTITLEVEPSDT
                     IENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIF
                     VKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQ
                     KESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGV"
ORIGIN
```

```
  Query Match             88.5%;  Score 1940;  DB 8;  Length 2207;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2140;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
```

```
Db      228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287
Qy      289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
Db      288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347
Qy      349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
Db      348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407
Qy      409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
Db      408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467
Qy      469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
Db      468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527
Qy      529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
Db      528 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 587
Qy      589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
Db      588 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 647
Qy      649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
Db      648 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 707
Qy      709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
Db      708 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 767
Qy      769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
Db      768 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 827
Qy      829 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 888
Db      828 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 887
Qy      889 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 948
Db      888 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 947
Qy      949 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1008
Db      948 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1007
Qy     1009 GTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1068
Db     1008 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1067
Qy     1069 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 1128
Db     1068 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 1127
Qy     1129 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 1188
Db     1128 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 1187
Qy     1189 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1248
Db     1188 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1247
Qy     1249 CTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1308
Db     1248 CTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1307
Qy     1309 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1368
Db     1308 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1367
Qy     1369 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTA 1428
Db     1368 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTA 1427
Qy     1429 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGG 1488
Db     1428 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGG 1487
Qy     1489 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1548
Db     1488 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1547
Qy     1549 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1608
Db     1548 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1607
Qy     1609 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1668
Db     1608 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1667
Qy     1669 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1728
Db     1668 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1727
Qy     1729 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1788
Db     1728 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1787
Qy     1789 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1848
Db     1788 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1847
Qy     1849 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 1908
Db     1848 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 1907
Qy     1909 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 1968
Db     1908 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 1967
Qy     1969 TCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTG 2028
Db     1968 TCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTG 2027
Qy     2029 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 2088
Db     2028 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 2087
Qy     2089 CTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAA 2148
Db     2088 CTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAA 2147
Qy     2149 CAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db     2148 CAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2191
```

```
RESULT 13
AK129749
LOCUS       AK129749               2191 bp    mRNA    linear   PRI 10-SEP-2003
DEFINITION  Homo sapiens cDNA FLJ26238 fis, clone CBR05440, highly similar to
            Rattus norvegicus ubiquitin C (Ubc).
ACCESSION   AK129749
VERSION     AK129749.1  GI:34526359
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M., Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2191)
AUTHORS Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .2191
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBR05440"
/tissue_type="brain"
/clone_lib="CBR"
/note="cloning vector: pME18SFL3"
ORIGIN

Query Match 87.0%; Score 1906; DB 8; Length 2191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 90

Qy      92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 150

Qy     152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 210

Qy     212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     211 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 270

Qy     272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 330

Qy     332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     331 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 390

Qy     392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     391 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 450

Qy     452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     451 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 510

Qy     512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     511 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 570

Qy     572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     571 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 630

Qy     632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     631 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 690

Qy     692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     691 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 750

Qy     752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     751 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 810

Qy     812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 871
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     811 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 870

Qy     872 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 931
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     871 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 930

Qy     932 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 991
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     931 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 990

Qy     992 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGA 1051
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     991 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1050

Qy    1052 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1111
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1051 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1110

Qy    1112 CTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGA 1171
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1111 CTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGA 1170

Qy    1172 GTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1231
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1171 GTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1230

Qy    1232 GACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1291
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1231 GACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1290

Qy    1292 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 1351
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1291 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1350

Qy    1352 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1351 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1410

Qy    1412 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1471
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1411 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1470

Qy    1472 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1531
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1471 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1530

Qy    1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1531 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1590

Qy    1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1651
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1591 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1650
```

```
Qy       1652 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1710

Qy       1712 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1771
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1711 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATTCAAGACAAGGAAGGCAT 1770

Qy       1772 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1830

Qy       1832 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 1891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 1890

Qy       1892 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 1951
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 GATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 1950

Qy       1952 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCA 2011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCA 2010

Qy       2012 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 2071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 2070

Qy       2072 CATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTC 2131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2071 CATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTC 2130

Qy       2132 CCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCC 2191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2131 CCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCC 2190

Qy       2192 C 2192
             |
Db       2191 C 2191
```

RESULT 14
AC126309/c
LOCUS AC126309 121984 bp DNA linear PRI 04-APR-2003
DEFINITION Homo sapiens 12 BAC RP11-592O2 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC126309
VERSION AC126309.8 GI:22003948
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121984)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 6 (bases 1 to 121984)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jul 30, 2002 this sequence version replaced gi:21913328.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES	Location/Qualifiers
source	1. .121984
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-592O2"
misc_feature	complement(1. .2001)
	/note="overlaps bases 1. .2001 of clone AC073593"
	/function="clone overlap"
repeat_region	319. .510
	/rpt_family="L2"
repeat_region	511. .637
	/rpt_family="FLAM_C"
repeat_region	638. .1013
	/rpt_family="L2"
repeat_region	1191. .1351
	/rpt_family="Charlie5"
repeat_region	1353. .1625
	/rpt_family="L1MC4"
STS	1583. .1866
	/standard_name="G59354"
repeat_region	1641. .1772
	/rpt_family="FLAM_C"
STS	2081. .2379
	/standard_name="SHGC-84469"
repeat_region	2829. .2869
	/rpt_family="MIR"
repeat_region	complement(3610. .3910)
	/rpt_family="AluSq"
repeat_region	4545. .4911
	/rpt_family="MLT1I"
repeat_region	5210. .5291
	/rpt_family="MIR"
repeat_region	5369. .5609
	/rpt_family="MLT1J"
repeat_region	5626. .5828
	/rpt_family="MLT1J"
repeat_region	6299. .6421
	/rpt_family="MLT1K"
repeat_region	7696. .7834
	/rpt_family="MIR"
repeat_region	7869. .7920
	/rpt_family="L2"
repeat_region	7921. .8217
	/rpt_family="AluSx"
repeat_region	8218. .8409
	/rpt_family="L2"
repeat_region	8534. .8692
	/rpt_family="L2"
repeat_region	8711. .8801
	/rpt_family="L2"
repeat_region	complement(9119. .9205)
	/rpt_family="MIR"
repeat_region	9358. .9504
	/rpt_family="(TG)n"
repeat_region	9585. .9716
	/rpt_family="FLAM_C"
repeat_region	complement(10426. .10580)
	/rpt_family="L2"
repeat_region	10671. .10797
	/rpt_family="AluJ"
repeat_region	complement(10799. .10968)
	/rpt_family="AluYa8"
repeat_region	complement(10971. .11291)
	/rpt_family="AluSx"
repeat_region	11292. .11493
	/rpt_family="AluJ"
repeat_region	12190. .12250
	/rpt_family="MIR"
repeat_region	complement(12256. .12380)
	/rpt_family="MER102"
STS	12488. .12761
	/standard_name="SHGC-111902"
repeat_region	complement(12595. .12666)
	/rpt_family="MIR"

Query Match 85.4%; Score 1873; DB 8; Length 121984;
Best Local Similarity .99.8%; Pred. No. 0;
Matches 2123; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106112 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 106053

Qy       125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    106052 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 105993

Qy       185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105992 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 105933

Qy       245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105932 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 105873

Qy       305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    105872 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCAT 105813

Qy       365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105812 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 105753

Qy       425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105752 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 105693

Qy       485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105692 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 105633

Qy       545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105632 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 105573

Qy       605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    105572 GGCAAAGATCCAAGATAAGGAAGGCATTCCTTCTGATCAGCAGAGGTTGATCTTTGCCGG 105513

Qy       665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105512 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 105453
```

```
Qy       725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
Db    105452 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 105393

Qy       785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
Db    105392 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 105333

Qy       845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
Db    105332 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 105273

Qy       905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT- 964
Db    105272 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 105213

Qy       965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
Db    105212 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 105153

Qy      1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
Db    105152 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 105093

Qy      1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
Db    105092 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 105033

Qy      1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
Db    105032 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 104973

Qy      1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
Db    104972 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 104913

Qy      1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
Db    104912 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 104853

Qy      1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
Db    104852 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 104793

Qy      1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
Db    104792 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 104733

Qy      1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
Db    104732 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 104673

Qy      1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
Db    104672 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 104613

Qy      1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
Db    104612 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 104553

Qy      1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
Db    104552 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 104493

Qy      1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
Db    104492 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 104433

Qy      1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
Db    104432 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 104373

Qy      1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
Db    104372 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 104313

Qy      1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1924
Db    104312 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAA 104253

Qy      1925 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1984
Db    104252 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 104193

Qy      1985 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2044
Db    104192 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 104133

Qy      2045 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2104
Db    104132 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 104073

Qy      2105 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2164
Db    104072 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 104013

Qy      2165 TTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db    104012 TTCCTTTCAATAAAGTTGTTGCATTCCC 103985
```

```
RESULT 15
AB089613
LOCUS       AB089613               2058 bp    DNA     linear   PRI 03-FEB-2004
DEFINITION  Homo sapiens UbC gene for polyubiquitin C, complete cds.
ACCESSION   AB089613
VERSION     AB089613.1  GI:28316335
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tachikui,H., Saitou,N., Nakajima,T., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Lineage-specific homogenization of the polyubiquitin gene among
            human and great apes
  JOURNAL   J. Mol. Evol. 57 (6), 737-744 (2003)
   PUBMED   14745543
REFERENCE   2  (bases 1 to 2058)
  AUTHORS   Tachikui,H., Nakajima,T., Saito,N., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2002) Hiroshi Tachikui, Institute of Medical
            Science, University of tokyo, Division of Genetic Dignosis;
            Shirokanedai 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:tachikui@ims.u-tokyo.ac.jp, Tel:81-5449-5750,
            Fax:81-5449-5764)
FEATURES             Location/Qualifiers
     source          1. .2058
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Japanese with 9 ub units"
     gene            1. .2058
                     /gene="UbC"
     CDS             1. .2058
                     /gene="UbC"
                     /codon_start=1
                     /product="polyubiquitin C"
                     /protein_id="BAC56951.1"
                     /db_xref="GI:28316336"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
```

TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
GGV"

ORIGIN

Query Match 84.6%; Score 1854; DB 8; Length 2058;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2054; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy    69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
Db     1 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 60

Qy   129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
Db    61 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 120

Qy   189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
Db   121 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 180

Qy   249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
Db   181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 240

Qy   309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
Db   241 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAG 300

Qy   369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
Db   301 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 360

Qy   429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
Db   361 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 420

Qy   489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
Db   421 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 480

Qy   549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
Db   481 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 540

Qy   609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
Db   541 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 600

Qy   669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
Db   601 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 660

Qy   729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
Db   661 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 720

Qy   789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
Db   721 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 780

Qy   849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
Db   781 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 840

Qy   909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
Db   841 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 900

Qy   969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028
Db   901 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 960

Qy  1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
Db   961 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1020

Qy  1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
Db  1021 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1080

Qy  1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
Db  1081 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1140

Qy  1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
Db  1141 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1200

Qy  1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
Db  1201 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1260

Qy  1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
Db  1261 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1320

Qy  1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448
Db  1321 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1380

Qy  1449 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1508
Db  1381 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1440

Qy  1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
Db  1441 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1500

Qy  1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
Db  1501 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1560

Qy  1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
Db  1561 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1620

Qy  1689 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1748
Db  1621 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1680

Qy  1749 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1808
Db  1681 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1740

Qy  1809 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1868
Db  1741 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1800

Qy  1869 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1928
Db  1801 CTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACC 1860

Qy  1929 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 1988
Db  1861 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 1920

Qy  1989 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 2048
```

```
Db       1921 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 1980

Qy       2049 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2040

Qy       2109 TTGAGGGGGGGTGTCTAA 2126
              ||||||||||||||||||
Db       2041 TTGAGGGGGGGTGTCTAA 2058


RESULT 16
HUMPOLYUBI
LOCUS       HUMPOLYUBI              6605 bp    DNA     linear   PRI 14-APR-2000
DEFINITION  Human UbC gene for polyubiquitin, exon1-2, partial cds.
ACCESSION   D63791
VERSION     D63791.1  GI:1304127
KEYWORDS    UbC; polyubiquitin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nenoi,M., Mita,K., Ichimura,S., Cartwright,I.L., Takahashi,E.,
            Yamauchi,M. and Tsuji,H.
  TITLE     Heterogeneous structure of the polyubiquitin gene UbC of HeLa S3
            cells
  JOURNAL   Gene 175 (1-2), 179-185 (1996)
   PUBMED   8917096
REFERENCE   2  (bases 1 to 6605)
  AUTHORS   Nenoi,M., Mita,K., Ichimura,S. and Cartwright,I.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 6605)
  AUTHORS   Nenoi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-1995) Mitsuru Nenoi, National Institute of
            Radiological Sciences, Division of Biology and Oncology; 9-1,
            Anagawa-4-chome, Inage-ku, Chiba, Chiba 263, Japan
            (E-mail:m_nenoi@nirs.go.jp, Tel:81-43-251-2111(ex.322),
            Fax:81-43-255-6497)
FEATURES             Location/Qualifiers
     source          1. .6605
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q24.31-32"
                     /cell_line="HeLa"
                     /clone_lib="library of T.Yamauchi"
     TATA_signal     3864. .3870
     prim_transcript 3895. .>6605
     exon            3895. .3958
                     /number=1
     intron          3959. .4769
                     /number=1
     exon            4770. .>6605
                     /number=2
     gene            4773. .6605
                     /gene="UbC"
     CDS             4773. .>6605
                     /gene="UbC"
                     /codon_start=1
                     /product="polyubiquitin"
                     /protein_id="BAA09860.1"
                     /db_xref="GI:1304128"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEGEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQI"
ORIGIN

  Query Match             74.5%;  Score 1633;  DB 8;  Length 6605;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1833;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4769 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 4828

Qy        125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4829 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 4888

Qy        185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4889 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 4948

Qy        245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4949 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 5008

Qy        305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5009 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 5068

Qy        365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5069 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 5128

Qy        425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5129 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 5188

Qy        485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5189 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 5248

Qy        545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5249 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 5308

Qy        605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5309 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 5368

Qy        665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5369 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 5428

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5429 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 5488

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5489 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 5548

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5549 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 5608

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5609 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 5668

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5669 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 5728
```

```
Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       5729 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAGGG 5788

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5789 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 5848

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5849 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 5908

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
             |||||||||||||||||||||||||||||||||||| |||||||||||||| ||||||
Db       5909 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACGATCACTCTCGAGGGGGAGCC 5968

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5969 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 6028

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6029 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 6088

Qy       1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6089 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 6148

Qy       1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6149 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 6208

Qy       1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6209 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 6268

Qy       1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6269 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 6328

Qy       1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6329 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 6388

Qy       1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6389 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 6448

Qy       1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6449 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 6508

Qy       1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6509 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 6568

Qy       1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1901
             |||||||||||||||||||||||||||||||||||||
Db       6569 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 6605
```

RESULT 17
AB003730
LOCUS AB003730 2291 bp DNA linear PRI 22-OCT-2002
DEFINITION Homo sapiens gene for polyubiquitin, complete cds.
ACCESSION AB003730
VERSION AB003730.1 GI:2627128
KEYWORDS polyubiquitin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nenoi,M., Mita,K., Ichimura,S. and Kawano,A.
TITLE Higher frequency of concerted evolutionary events in rodents than in man at the polyubiquitin gene VNTR locus
JOURNAL Genetics 148 (2), 867-876 (1998)
PUBMED 9504932
REFERENCE 2 (bases 1 to 2291)
AUTHORS Nenoi,M.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-1997) Mitsuru Nenoi, National Institute of Radiological Sciences, Division of Biology and Oncology; 9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263, Japan (E-mail:m_nenoi@nirs.go.jp, Tel:81-43-251-2111(ex.322), Fax:81-43-255-6497)
FEATURES Location/Qualifiers
source 1. .2291
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24"
/cell_line="HeLa"
gene 425. .2254
/gene="UbC"
CDS 425. .2254
/gene="UbC"
/codon_start=1
/product="polyubiquitin"
/protein_id="BAA23486.1"
/db_xref="GI:2627129"
/translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
LRLRGGMQIFVKTLTGKTITXEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGV"
ORIGIN

Query Match 58.6%; Score 1284; DB 8; Length 2291;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy         65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 480

Qy        125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 540

Qy        185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 600

Qy        245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 660

Qy        305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        661 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCAT 720

Qy        365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 780

Qy        425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
```

```
Db        781 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 840
Qy        485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
Db        841 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 900
Qy        545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
Db        901 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 960
Qy        605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
Db        961 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 1020
Qy        665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
Db       1021 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 1080
Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
Db       1081 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1140
Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
Db       1141 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1200
Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
Db       1201 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1260
Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
Db       1261 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1320
Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
Db       1321 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 1380
Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
Db       1381 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1440
Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
Db       1441 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1500
Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
Db       1501 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1560
Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
Db       1561 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1620
Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
Db       1621 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1680
Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
Db       1681 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1740
Qy       1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
Db       1741 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1800
Qy       1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
Db       1801 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCNCGAAGTGGAGCCGAGTGACACCAT 1860
Qy       1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
Db       1861 CGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1920
Qy       1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
Db       1921 GATCTTTGCCGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1980
Qy       1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
Db       1981 AGAGTCCACCCTGCACCTGGTGCTCCGTCT 2010
```

```
RESULT 18
AX337205
LOCUS       AX337205               2309 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 7714 from Patent WO0194629.
ACCESSION   AX337205
VERSION     AX337205.1  GI:18127924
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 7714 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           58.5%;  Score 1283;  DB 6;  Length 2309;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216
Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276
Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336
Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396
Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456
Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516
Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
Db        517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576
Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
Db        577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636
```

```
Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696

Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756

Qy       1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816

Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876

Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936

Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996

Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056

Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116

Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176

Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236

Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296

Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356

Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416

Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643


RESULT 19
HUMUBI13
LOCUS       HUMUBI13                2309 bp    mRNA    linear   PRI 03-AUG-1993
DEFINITION  Human ubiquitin mRNA, complete cds.
ACCESSION   M26880
VERSION     M26880.1  GI:340067
KEYWORDS    ubiquitin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2309)
  AUTHORS   Wiborg,O., Pedersen,M.S., Wind,A., Berglund,L.E., Marcker,K.A. and
            Vuust,J.
  TITLE     The human ubiquitin multigene family: some genes contain multiple
            directly repeated ubiquitin coding sequences
  JOURNAL   EMBO J. 4 (3), 755-759 (1985)
  PUBMED    2988935
COMMENT     Original source text: Human adult leukocyte, cDNA to mRNA, clone
            lambda-HUb13.
FEATURES             Location/Qualifiers
     source          1. .2309
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             136. .2193
                     /note="ubiquitin"
                     /codon_start=1
                     /protein_id="AAA36789.1"
                     /db_xref="GI:340068"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGV"
     unsure          1983
                     /note="g could be t"
     polyA_signal    2240. .2245
ORIGIN

  Query Match           58.5%;  Score 1283;  DB 8;  Length 2309;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              |||||||||||||||||| ||||||||||||||||| |||||||||||||| ||||||||||||
Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516
```

```
Qy       906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
Db       517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576

Qy       966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
Db       577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636

Qy      1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
Db       637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696

Qy      1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
Db       697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756

Qy      1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
Db       757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816

Qy      1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
Db       817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876

Qy      1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
Db       877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936

Qy      1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
Db       937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996

Qy      1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
Db       997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056

Qy      1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
Db      1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116

Qy      1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
Db      1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176

Qy      1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
Db      1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236

Qy      1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
Db      1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296

Qy      1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
Db      1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356

Qy      1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
Db      1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416

Qy      1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
Db      1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476

Qy      1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
Db      1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536

Qy      1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
Db      1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596

Qy      1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
Db      1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643
```

```
RESULT 20
AK097415
LOCUS       AK097415               1864 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ40096 fis, clone TESTI2004080, highly similar
            to Homo sapiens mRNA for polyubiquitin UbC.
ACCESSION   AK097415
VERSION     AK097415.1  GI:21757163
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,
            Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
            Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1864)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
```

HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES Location/Qualifiers
source 1. .1864
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI2004080"
/tissue_type="testis"
/clone_lib="TESTI2"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 52.6%; Score 1154; DB 8; Length 1864;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 307

Qy       408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 367

Qy       468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 427

Qy       528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       428 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 487

Qy       588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       488 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 547

Qy       648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       548 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 607

Qy       708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       608 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 667

Qy       768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       668 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 727

Qy       828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       728 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 787

Qy       888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       788 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 847

Qy       948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       848 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 907

Qy      1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       908 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 967

Qy      1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       968 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1027

Qy      1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1028 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1087

Qy      1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1088 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1147

Qy      1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1148 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1207

Qy      1308 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1367
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      1208 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1267

Qy      1368 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1427
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1268 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1327

Qy      1428 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1487
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1328 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1387

Qy      1488 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1547
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1388 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1447

Qy      1548 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 1607
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1448 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 1507

Qy      1608 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1508 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1554
```

RESULT 21
BC008955
LOCUS BC008955 1728 bp mRNA linear PRI 04-NOV-2003
DEFINITION Homo sapiens ubiquitin C, mRNA (cDNA clone IMAGE:3030313), partial cds.
ACCESSION BC008955
VERSION BC008955.2 GI:38114660
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1728)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1728)
AUTHORS Strausberg,R.
TITLE Direct Submission

```
  JOURNAL   Submitted (29-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     On Oct 31, 2003 this sequence version replaced gi:14286307.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 10 Row: b Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 20149305.
FEATURES             Location/Qualifiers
     source          1. .1728
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3030313"
                     /tissue_type="Ovary, adenocarcinoma"
                     /clone_lib="NIH_MGC_9"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            <1. .1728
                     /gene="UBC"
                     /note="synonym: HMG20"
                     /db_xref="GeneID:7316"
                     /db_xref="MIM:191340"
     CDS             <1. .1643
                     /gene="UBC"
                     /codon_start=3
                     /product="UBC protein"
                     /protein_id="AAH08955.2"
                     /db_xref="GI:38114661"
                     /db_xref="GeneID:7316"
                     /db_xref="MIM:191340"
                     /translation="QFLFVDCCDRHLTMQIFVKTLTGKTITLEVEPSDTIENVKAKIQ
                     DKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTI
                     TLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVL
                     RLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDG
                     RTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEG
                     IPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEV
                     EPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRG
                     GMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLS
                     DYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPD
                     QQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGV"
     misc_feature    42. .263
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    270. .491
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    498. .719
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    726. .947
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    954. .1175
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    1182. .1403
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    1410. .1631
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
ORIGIN

  Query Match             51.2%;  Score 1123;  DB 8;  Length 1728;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1373;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 81

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 141

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 201

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 261

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        262 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 321

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 381

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 441

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 501

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 561

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 621

Qy        649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 681
```

```
Qy        709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 741

Qy        769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 801

Qy        829 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 861

Qy        889 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        862 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 921

Qy        949 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 981

Qy       1009 GTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1068
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1041

Qy       1069 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 1101

Qy       1129 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 1188
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1102 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1161

Qy       1189 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1221

Qy       1249 CTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1308
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 CTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1281

Qy       1309 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1368
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1282 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1341

Qy       1369 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1399
```

RESULT 22
BC014880
LOCUS BC014880 1122 bp mRNA linear PRI 30-SEP-2003
DEFINITION Homo sapiens ubiquitin C, mRNA (cDNA clone IMAGE:3907070), complete cds.
ACCESSION BC014880
VERSION BC014880.1 GI:15928839
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1122)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1122)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 15 Row: b Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24496788
This clone has the following problem: The cds is short compared to the longest cds in the locus.

FEATURES Location/Qualifiers
source 1. .1122
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3907070"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
gene 1. .1122
/gene="UBC"
/note="synonym: HMG20"
/db_xref="GeneID:7316"
/db_xref="MIM:191340"
CDS 120. .1037
/gene="UBC"
/codon_start=1
/product="UBC protein"
/protein_id="AAH14880.1"
/db_xref="GI:15928840"
/db_xref="GeneID:7316"
/db_xref="MIM:191340"
/translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK

```
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGV"
     misc_feature    120. .341
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    348. .569
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    576. .797
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
     misc_feature    804. .1025
                     /gene="UBC"
                     /note="ubiquitin; Region: Ubiquitin family. This family
                     contains a number of ubiquitin-like proteins: SUMO (smt3
                     homologue), Nedd8, Elongin B, Rub1"
                     /db_xref="CDD:pfam00240"
ORIGIN

  Query Match            50.3%;  Score 1103;  DB 8;  Length 1122;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1103;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1090 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGT 1149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGT 60

Qy       1150 CTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 1209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 120

Qy       1210 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTG 1269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTG 180

Qy       1270 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGC 1329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGC 240

Qy       1330 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 1389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 300

Qy       1390 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 1449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCG 360

Qy       1450 TGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGA 1509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGA 420

Qy       1510 ATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCT 1569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCT 480

Qy       1570 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 1629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 540

Qy       1630 CCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 1689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 600

Qy       1690 CTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAA 1749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAA 660

Qy       1750 AGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAAC 1809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAAC 720

Qy       1810 AGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACC 1869
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACC 780

Qy       1870 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 1929
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 840

Qy       1930 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 1989
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 900

Qy       1990 AGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 2049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 960

Qy       2050 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCT 2109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCT 1020

Qy       2110 TGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCT 2169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCT 1080

Qy       2170 TTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||
Db       1081 TTCAATAAAGTTGTTGCATTCCC 1103


RESULT 23
AK092983
LOCUS       AK092983               2784 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ35664 fis, clone SPLEN2017687, highly similar
            to Polyubiquitin 9.
ACCESSION   AK092983
VERSION     AK092983.1  GI:21751700
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
```

```
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
            Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
            Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
            Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
            Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2784)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .2784
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="SPLEN2017687"
                     /tissue_type="spleen"
                     /clone_lib="SPLEN2"
                     /note="cloning vector: pME18SFL3"
ORIGIN

  Query Match             46.4%;  Score 1017;  DB 8;  Length 2784;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1167;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1615 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1674

Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1675 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1734

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1735 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1794

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1795 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGGCCATCACTCTCGAGGTGGAG 1854

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1855 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1914

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1915 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1974

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1975 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 2034

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2035 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 2094

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2095 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 2154

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2155 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2214

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2215 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 2274

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2275 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 2334

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2335 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 2394

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 2454

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       2455 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGT 2514

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 2574

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2634

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2694

Qy       2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2695 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2754

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||||||
Db       2755 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2784


RESULT 24
AB209782
LOCUS       AB209782               3996 bp    mRNA    linear   PRI 31-MAR-2005
DEFINITION  Homo sapiens mRNA for ubiquitin C variant protein.
ACCESSION   AB209782
VERSION     AB209782.1  GI:62089149
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
            Ohara,O., Nagase,T. and F.Kikuno,R.
  TITLE     None Title
```

```
  JOURNAL   Published Only in Database (2005)
REFERENCE   2  (bases 1 to 3996)
  AUTHORS   Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A., Yokoyama,S.,
            Ohara,O., Nagase,T. and F.Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,
            Department of Human Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
            Chiba, 292-0818, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://protein.gsc.riken.go.jp/, Tel:81-438-52-3930,
            Fax:81-438-52-3931)
COMMENT     This work was supported in part by the National Project on Protein
            Structural and Functional Analysis , Ministry of Education, Culture
            , Sports, Science and Technology of Japan. Totoki Y, Toyoda A,
            Takeda T, Sakaki Y, Tanaka A, Yokoyama S. RIKEN Genomic Sciences
            Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
            e-mail: aktanaka@postman.riken.go.jp
            URL: http://protein.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .3996
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="bm02564"
                     /tissue_type="brain"
                     /note="this clone is also named as hsk003002240
                     vector:pBluescriptII SK plus"
     gene            <1. .3996
                     /gene="ubiquitin C variant"
     CDS             <1. .3930
                     /gene="ubiquitin C variant"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /evidence=not_experimental
                     /product="ubiquitin C variant"
                     /protein_id="BAD93019.1"
                     /db_xref="GI:62089150"
                     /translation="FGSQFLFVDRCDRHLTMQIFVKTLTGKTITLEVEPSDTIENVKA
                     KIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTG
                     KTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLH
                     LVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQL
                     EDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQD
                     KEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTIT
                     LEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLR
                     LRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGR
                     TLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGI
                     PPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVE
                     PSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
                     MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSD
                     YNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQ
                     QRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDT
                     IENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIF
                     VKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQ
                     KESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLENGRTLSDYNIQKESTLHLVLRLRGGV"
ORIGIN

  Query Match           38.2%;  Score 837;  DB 8;  Length 3996;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches  987;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 71

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 131

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 191

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 251

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 311

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        312 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 371

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 431

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 491

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 551

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 611

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 671

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 731

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 791

Qy        812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 871
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 TGACACCATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 851

Qy        872 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 911

Qy        932 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 991
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        912 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 971

Qy        992 CGTGAAGACCCTGACTGGTAAGACCATCAC 1021
              |||||||||||||||||||||||||||||
Db        972 CGTGAAGACCCTGACTGGTAAGACCATCAC 1001


RESULT 25
AF052164
LOCUS       AF052164               1168 bp    mRNA    linear   PRI 05-AUG-1998
DEFINITION  Homo sapiens clone 24683 mRNA sequence.
ACCESSION   AF052164
VERSION     AF052164.1  GI:3360475
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Andersson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.
  TITLE     A 'double adaptor' method for improved shotgun library construction
  JOURNAL   Anal. Biochem. 236 (1), 107-113 (1996)
   PUBMED   8619474
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Yu,W., Andersson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
            Ricafrente,J.Y., Wentland,M.A., Lennon,G. and Gibbs,R.A.
  TITLE     Large-scale concatenation cDNA sequencing
  JOURNAL   Genome Res. 7 (4), 353-358 (1997)
   PUBMED   9110174
REFERENCE   3  (bases 1 to 1168)
  AUTHORS   Yu,W., Sarginson,J. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAR-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="I.M.A.G.E. Consortium clone ID 24683"
                     /sex="female"
                     /tissue_type="brain"
                     /clone_lib="1NIB"
                     /dev_stage="infant"
                     /note="similar to Homo sapiens polyubiquitin encoded by
                     GenBank Accession Number AB003730"
ORIGIN

  Query Match             37.7%;  Score 827;  DB 8;  Length 1168;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  827;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 60

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 120

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 180

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 240

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 300

Qy       1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 360

Qy       1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 420

Qy       1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 480

Qy       1308 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 540

Qy       1368 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 600

Qy       1428 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 660

Qy       1488 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 720

Qy       1548 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 1607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCT 780

Qy       1608 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 827


RESULT 26
HUMUB
LOCUS       HUMUB                     824 bp    mRNA    linear   PRI 03-AUG-1993
DEFINITION  Human poly-ubiquitin mRNA, complete cds.
ACCESSION   M17597
VERSION     M17597.1  GI:340057
KEYWORDS    poly-ubiquitin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 824)
  AUTHORS   Einspanier,R., Sharma,H.S. and Scheit,K.H.
  TITLE     Cloning and sequence analysis of a cDNA encoding poly-ubiquitin in
            human ovarian granulosa cells
  JOURNAL   Biochem. Biophys. Res. Commun. 147 (2), 581-587 (1987)
   PUBMED   2820408
COMMENT     Original source text: Human ovarian granulosa cell, cDNA to mRNA,
            clone pHGR21.
            Draft enty and computer-readable sequence for [1] kindly provided
            by R.Einspanier and K.H.Scheit, 04-DEC-1987.
            In human ovarian granulosa cells two main mRNA species were
            detected differing in size and amount.  The reported poly-ubiquitin
            cDNA sequence reported below is not homologous to the known
            ubiquitin-9 repeat or to the 3-repeat gene.
FEATURES             Location/Qualifiers
     source          1. .824
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     mRNA            <1. .>824
                     /product="UB mRNA"
     CDS             <1. .810
                     /note="ubiquitin precursor"
                     /codon_start=1
                     /protein_id="AAA36787.1"
                     /db_xref="GI:340058"
                     /translation="PPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIF
                     VKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQ
                     KESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGV"
     mat_peptide     <1. .120
                     /product="ubiquitin copy A"
     mat_peptide     121. .348
                     /product="ubiquitin copy B"
     mat_peptide     349. .576
                     /product="ubiquitin copy C"
     mat_peptide     577. .807
                     /product="ubiquitin copy D"
ORIGIN      6 bp upstream of Sau3A site.

  Query Match             37.6%;  Score 824;  DB 8;  Length 824;
  Best Local Similarity   100.0%;  Pred. No. 0;
```

```
  Matches  824;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1317 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 60

Qy       1377 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 120

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 180

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 240

Qy       1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 300

Qy       1617 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 360

Qy       1677 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 420

Qy       1737 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 480

Qy       1797 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 540

Qy       1857 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 600

Qy       1917 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 1976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 660

Qy       1977 AAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAA 2036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAA 720

Qy       2037 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCAC 2096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCAC 780

Qy       2097 TTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAA 2140
             ||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAA 824


RESULT 27
HSM801035
LOCUS       HSM801035               796 bp    mRNA    linear   PRI 16-APR-2005
DEFINITION  Homo sapiens mRNA; cDNA DKFZp434K0435 (from clone DKFZp434K0435);
            partial cds.
ACCESSION   AL117514
VERSION     AL117514.1  GI:5912027
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) MIPS, Am Klopferspitz 18a, D-82152
            Martinsried, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp434K0435) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp434K0435"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434K0435"
                     /tissue_type="testis"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     gene            <1. .796
                     /gene="DKFZp434K0435"
     CDS             <1. .720
                     /gene="DKFZp434K0435"
                     /note="polyubiquitin 9"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB55973.1"
                     /db_xref="GI:5912028"
                     /db_xref="GOA:Q9UFQ0"
                     /db_xref="InterPro:IPR000626"
                     /db_xref="UniProt/TrEMBL:Q9UFQ0"
                     /translation="LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL
                     SDYNIQKESTLHLVLRLRGGV"
     polyA_signal    767. .772
                     /gene="DKFZp434K0435"
     polyA_site      786
                     /gene="DKFZp434K0435"
ORIGIN

  Query Match             33.5%;  Score 735;  DB 8;  Length 796;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  785;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1407 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 60

Qy       1467 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1526
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 120

Qy       1527 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 1586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 180

Qy       1587 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 240

Qy       1647 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1706
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 300

Qy       1707 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1766
```

```
Db        301 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 360

Qy       1767 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1826

Db        361 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 420

Qy       1827 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 1886

Db        421 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 480

Qy       1887 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 1946

Db        481 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 540

Qy       1947 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 2006

Db        541 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 600

Qy       2007 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 2066

Db        601 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 660

Qy       2067 TACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAA 2126

Db        661 TACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAA 720

Qy       2127 GTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGC 2186

Db        721 GTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGC 780

Qy       2187 ATTCCC 2192

Db        781 ATTCCC 786
```

```
RESULT 28
AX714536
LOCUS       AX714536               1300 bp    DNA     linear   PAT 15-APR-2003
DEFINITION  Sequence 1220 from Patent EP1293569.
ACCESSION   AX714536
VERSION     AX714536.1  GI:29889489
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuho,Y.
  TITLE     Full-length cDNAs
  JOURNAL   Patent: EP 1293569-A 1220 19-MAR-2003;
            Helix Research Institute (JP) ; Research Association for
            Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1300
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             32.8%;  Score 719;  DB 6;  Length 1300;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  769;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407

Db        257 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 316

Qy        408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467

Db        317 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 376

Qy        468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527

Db        377 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 436

Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587

Db        437 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 496

Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647

Db        497 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 556

Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707

Db        557 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 616

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767

Db        617 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 676

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827

Db        677 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 736

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887

Db        737 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 796

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947

Db        797 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 856

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007

Db        857 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 916

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067

Db        917 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 976

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1117

Db        977 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1026
```

```
RESULT 29
AK056939
LOCUS       AK056939               1300 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ32377 fis, clone SKMUS1000014, highly similar
            to Polyubiquitin 9.
ACCESSION   AK056939
VERSION     AK056939.1  GI:16552474
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
```

Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs

JOURNAL Nat. Genet. 36 (1), 40-45 (2004)

PUBMED 14702039

REFERENCE 2

AUTHORS Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K., Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1300)

AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.

TITLE Direct Submission

JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES Location/Qualifiers

source 1. .1300
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SKMUS1000014"
/tissue_type="skeletal muscle"
/clone_lib="SKMUS1"
/note="cloning vector: pME18SFL3"

CDS 68. .1234
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71316.1"
/db_xref="GI:16552475"
/translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGV"

ORIGIN

Query Match 32.8%; Score 719; DB 8; Length 1300;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       257 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 316

Qy       408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       317 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 376

Qy       468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       377 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 436

Qy       528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       437 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 496

Qy       588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       497 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 556

Qy       648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       557 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 616

Qy       708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       617 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 676

Qy       768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       677 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 736

Qy       828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       737 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 796

Qy       888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       797 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 856

Qy       948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       857 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 916

Qy      1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       917 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 976

Qy      1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1117
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       977 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1026
```

RESULT 30
BV177540
LOCUS BV177540 1201 bp DNA linear STS 10-JUN-2004
DEFINITION sqnm95284 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
ACCESSION BV177540
VERSION BV177540.1 GI:48013657
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions
JOURNAL Genome Res. (2004) In press

```
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1201.
FEATURES              Location/Qualifiers
     source           1. .1201
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>1201
ORIGIN

  Query Match            32.7%;  Score 716;  DB 10;  Length 1201;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches  936;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        311 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 66

Qy        371 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 126

Qy        431 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 186

Qy        491 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 550
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        187 TACCCTGCACCTGGT-CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 245

Qy        551 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 305

Qy        611 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 365

Qy        671 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 425

Qy        731 GGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 485

Qy        791 CACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAA 850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAA 545

Qy        851 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGG 910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGG 605

Qy        911 GCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCT 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 GCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCT 665

Qy        971 CAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGT 1030
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        666 CAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACYCTCGAAGT 725

Qy       1031 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 785

Qy       1091 TCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTC 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 TCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTC 845

Qy       1151 TGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1210
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        846 TGACTACAACATCCAGAAAGAGTCCACCNTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 905

Qy       1211 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1249
              |||||||||||||||||||||||||||||||||||||||
Db        906 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 944


RESULT 31
BV176676/c
LOCUS       BV176676               1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm90309 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV176676
VERSION     BV176676.1  GI:48012564
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES              Location/Qualifiers
     source           1. .1251
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>1251
ORIGIN

  Query Match            31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              ||||||||||||||||||||||
Db        580 CAACATCCAGAAAGAGTCCACC 559
```

```
RESULT 32
BV176677/c
LOCUS       BV176677               1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm90313 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV176677
VERSION     BV176677.1  GI:48012565
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES             Location/Qualifiers
     source          1. .1251
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1251
ORIGIN
```

```
  Query Match           31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              ||||||||||||||||||||||
Db        580 CAACATCCAGAAAGAGTCCACC 559
```

```
RESULT 33
BV176678/c
LOCUS       BV176678               1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm90314 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV176678
VERSION     BV176678.1  GI:48012567
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
```

```
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES              Location/Qualifiers
     source           1. .1251
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>1251
ORIGIN

  Query Match           31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              ||||||||||||||||||||||
Db        580 CAACATCCAGAAAGAGTCCACC 559
```

```
RESULT 34
BV177180/c
LOCUS       BV177180                1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm93250 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV177180
VERSION     BV177180.1  GI:48013200
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES              Location/Qualifiers
     source           1. .1251
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone_lib="Human DNA (Sequenom)"
     STS              <1. .>1251
ORIGIN

  Query Match           31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              |||||||||||||||||||||||
Db        580 CAACATCCAGAAAGAGTCCACC 559


RESULT 35
BV177599/c
LOCUS       BV177599               1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm95667 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV177599
VERSION     BV177599.1  GI:48013731
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES             Location/Qualifiers
     source          1. .1251
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1251
ORIGIN

  Query Match             31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              |||||||||||||||||||||||
Db        580 CAACATCCAGAAAGAGTCCACC 559


RESULT 36
BV180128/c
LOCUS       BV180128               1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm108117 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV180128
VERSION     BV180128.1  GI:48017094
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES             Location/Qualifiers
     source          1. .1251
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1251
ORIGIN

  Query Match           31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              |||||||||||||||||||||||
Db         580 CAACATCCAGAAAGAGTCCACC 559


RESULT 37
BV180129/c
LOCUS       BV180129                1251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm108121 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV180129
VERSION     BV180129.1  GI:48017096
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1251.
FEATURES             Location/Qualifiers
     source          1. .1251
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1251
ORIGIN

  Query Match           31.1%;  Score 682;  DB 10;  Length 1251;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  682;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1181

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1121

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1061

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1001

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 941

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 881

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 821

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         820 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 761

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         760 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 701

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         700 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 641

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         640 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 581

Qy       1385 CAACATCCAGAAAGAGTCCACC 1406
              |||||||||||||||||||||||
```

```
Db        580 CAACATCCAGAAAGAGTCCACC 559


RESULT 38
AY819759
LOCUS       AY819759                718 bp    mRNA    linear   PRI 28-NOV-2004
DEFINITION  Homo sapiens ubiquitin C splice variant (UBC) mRNA, complete cds,
            alternatively spliced.
ACCESSION   AY819759
VERSION     AY819759.1  GI:55977014
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Wang,P.Z., Wang,X. and Wu,J.
  TITLE     Novel splicing variants of some human genes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 718)
  AUTHORS   Wang,P.Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2004) Chinese National Human Genome Center, 3#
            707 Yongchang Road, BDA, Beijing 100176, P.R.China
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .718
                     /gene="UBC"
     CDS             19. .708
                     /gene="UBC"
                     /note="alternatively spliced"
                     /codon_start=1
                     /product="ubiquitin C splice variant"
                     /protein_id="AAV68344.1"
                     /db_xref="GI:55977015"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGV"
ORIGIN

  Query Match             28.8%;  Score 632;  DB 8;  Length 718;
  Best Local Similarity   100.0%;  Pred. No: 0;
  Matches  632;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          51 TGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACC 110
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACC 60

Qy         111 CTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 120

Qy         171 GGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGC 230
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGC 180

Qy         231 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 290
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 240

Qy         291 GGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAG 350
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAG 300

Qy         351 CCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCT 410
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCT 360

Qy         411 GACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGAC 470
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGAC 420

Qy         471 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 530
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 480

Qy         531 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 590
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 540

Qy         591 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGG 650
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGG 600

Qy         651 TTGATCTTTGCCGGAAAACAGCTGGAAGATGG 682
               ||||||||||||||||||||||||||||||||
Db         601 TTGATCTTTGCCGGAAAACAGCTGGAAGATGG 632


RESULT 39
AK026846
LOCUS       AK026846                1082 bp    mRNA    linear   PRI 12-SEP-2003
DEFINITION  Homo sapiens cDNA: FLJ23193 fis, clone REC00306, highly similar to
            AF052164 Homo sapiens clone 24683 mRNA sequence.
ACCESSION   AK026846
VERSION     AK026846.1  GI:10439801
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
            Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
            Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
            Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
            Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1082)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .1082
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="REC00306"
                     /cell_type="primary epithelial cells of human renal
                     proximal tubule"
                     /clone_lib="REC"
                     /note="cloning vector pME18SFL3"
     misc_feature    1. .1082
                     /note="highly similar to AF052164 Homo sapiens clone 24683
                     mRNA sequence"
ORIGIN
```

```
  Query Match           26.0%;  Score 571;  DB 8;  Length 1082;
  Best Local Similarity  100.0%;  Pred. No. 4.7e-308;
  Matches  571;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 71

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 131

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 191

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 251

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 311

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 371

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 431

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 491

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 551

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||
Db        552 AGACCCTGACTGGTAAGACCATCACCCTCGA 582


RESULT 40
AB089614
LOCUS       AB089614               2286 bp    DNA     linear   PRI 03-FEB-2004
DEFINITION  Pan troglodytes UbC gene for polyubiquitin C, complete cds.
ACCESSION   AB089614
VERSION     AB089614.1  GI:28316337
KEYWORDS    .
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Tachikui,H., Saitou,N., Nakajima,T., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Lineage-specific homogenization of the polyubiquitin gene among
            human and great apes
  JOURNAL   J. Mol. Evol. 57 (6), 737-744 (2003)
   PUBMED   14745543
REFERENCE   2  (bases 1 to 2286)
  AUTHORS   Tachikui,H., Nakajima,T., Saito,N., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2002) Hiroshi Tachikui, Institute of Medical
            Science, University of tokyo, Division of Genetic Dignosis;
            Shirokanedai 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:tachikui@ims.u-tokyo.ac.jp, Tel:81-5449-5750,
            Fax:81-5449-5764)
FEATURES             Location/Qualifiers
     source          1. .2286
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /note="Chimpanzee with 10 ub units"
     gene            1. .2286
                     /gene="UbC"
     CDS             1. .2286
                     /gene="UbC"
                     /codon_start=1
                     /product="polyubiquitin C"
                     /protein_id="BAC56952.1"
                     /db_xref="GI:28316338"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL
                     SDYNIQKESTLHLVLRLRGGV"
ORIGIN

  Query Match           25.4%;  Score 556;  DB 8;  Length 2286;
  Best Local Similarity  99.5%;  Pred. No. 1.2e-299;
  Matches  756;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        656 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 715

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        716 AGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 775

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        776 AAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 835

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        836 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 895

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        896 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 955

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        956 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1015

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1016 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1075

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1076 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1135

Qy        748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1136 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 1195

Qy        808 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 867
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1196 CCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 1255

Qy        868 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1256 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 1315

Qy        928 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1316 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1375

Qy        988 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1027
              ||||||||||||||||||||||||||||||||||||||||
Db       1376 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1415


RESULT 41
BV177342
LOCUS       BV177342               1110 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm94123 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV177342
VERSION     BV177342.1  GI:48013404
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1110.
FEATURES             Location/Qualifiers
     source          1. .1110
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1110
ORIGIN

  Query Match             25.3%;  Score 554;  DB 10;  Length 1110;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-298;
  Matches  704;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         126 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 185

Qy         126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 245

Qy         186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 305

Qy         246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 365

Qy         306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
               |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         366 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTNGAGCCCAGTGACACCATC 425

Qy         366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 485

Qy         426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
               |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         486 ATCTTTGCCGGAAAGCAGCTGGAAGATGGNCGCACCCTGTCTGACTACAACATCCAGAAA 545

Qy         486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
               ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         546 GAGTCNACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 605

Qy         546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         606 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 665

Qy         606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         666 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 725

Qy         666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         726 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 785

Qy         726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 772
               |||||||||||||||||||||||||||||||||||||||||||||||
Db         786 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 832


RESULT 42
BV179342
LOCUS       BV179342               1110 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm104938 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV179342
VERSION     BV179342.1  GI:48016048
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1110.
FEATURES             Location/Qualifiers
     source          1. .1110
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1110
ORIGIN
```

```
  Query Match           25.3%;  Score 554;  DB 10;  Length 1110;
  Best Local Similarity  99.6%;  Pred. No. 1.6e-298;
  Matches  704;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 185

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 245

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 305

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 365

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        366 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTNGAGCCCAGTGACACCATC 425

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 485

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        486 ATCTTTGCCGGAAAGCAGCTGGAAGATGGNCGCACCCTGTCTGACTACAACATCCAGAAA 545

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GAGTCNACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 605

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 665

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 725

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 785

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 772
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        786 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 832


RESULT 43
BV176674/c
LOCUS       BV176674               1063 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm90307 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV176674
VERSION     BV176674.1  GI:48012562
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1063.
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1063
ORIGIN

  Query Match           20.8%;  Score 456;  DB 10;  Length 1063;
  Best Local Similarity  99.0%;  Pred. No. 1.4e-243;
  Matches  906;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1278 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        936 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 877

Qy       1338 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        876 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 817

Qy       1398 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        816 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 757

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        756 CTGACTGGTAAGACCATCACTCTCGARGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 697

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 637

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 577

Qy       1638 CACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1697
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        576 CACCTGGTGCTCCGTCTTAGAGGTNGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 517

Qy       1698 ACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 NCCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 457

Qy       1758 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 1817
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGNTGATCTTTGCTGGGAAACAGCTGGAA 397

Qy       1818 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 337

Qy       1878 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1937
              | ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        336 CNTCTCAGAGGTGGGATGCANATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 277

Qy       1938 GAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 1997
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GANGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 217

Qy       1998 ATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        216 ATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 157

Qy       2058 CTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGG 2117
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CTGTCNGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGG 97

Qy       2118 GGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAA 2177
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GGTGTNTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAA 37

Qy       2178 AGTTGTTGCATTCCC 2192
              ||||||||||||||||
Db         36 AGTTGTTGCATTCCC 22


RESULT 44
BV176675/c
LOCUS       BV176675                1063 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm90308 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV176675
VERSION     BV176675.1  GI:48012563
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 1063.
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>1063
ORIGIN

  Query Match          20.8%;  Score 456;  DB 10;  Length 1063;
  Best Local Similarity   99.0%;  Pred. No. 1.4e-243;
  Matches  906;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1278 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        936 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 877

Qy       1338 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        876 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 817

Qy       1398 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        816 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 757

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        756 CTGACTGGTAAGACCATCACTCTCGANGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 697

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 637

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 577

Qy       1638 CACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1697
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        576 CACCTGGTGCTCCGTCTTAGAGGTNGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 517

Qy       1698 ACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 NCCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 457

Qy       1758 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 1817
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        456 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGNTGATCTTTGCTGGGAAACAGCTGGAA 397

Qy       1818 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 337

Qy       1878 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1937
              | ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        336 CNTCTCAGAGGTGGGATGCARATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 277

Qy       1938 GAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 1997
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GANGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 217

Qy       1998 ATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 157

Qy       2058 CTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGG 2117
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CTGTCNGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGG 97

Qy       2118 GGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAA 2177
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GGTGTNTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAA 37

Qy       2178 AGTTGTTGCATTCCC 2192
              ||||||||||||||||
Db         36 AGTTGTTGCATTCCC 22


RESULT 45
BD237119
LOCUS       BD237119                547 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compounds for remedy and diagnosis of lung cancer and method for
            using the same.
ACCESSION   BD237119
VERSION     BD237119.1  GI:33046889
KEYWORDS    JP 2002516659-A/120.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Reed,S.G., Lodes,M.J., Frudakis,T.N. and Mohamath,R.
  TITLE     Compounds for remedy and diagnosis of lung cancer and method for
            using the same
  JOURNAL   Patent: JP 2002516659-A 120 11-JUN-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002516659-A/120
            PD   11-JUN-2002
            PF   26-JAN-1999 JP 2000529432
```

```
            PR   28-JAN-1998 US    09/015029,28-JAN-1998 US    09/015022 PR
            18-MAR-1998 US    09/040828,18-MAR-1998 US    09/040831 PR
            23-JUL-1998 US    09/122192,23-JUL-1998 US    09/122191 PR
            22-DEC-1998 US    09/219245
            PI   STEVEN G REED,MICHAEL J LODES,TONY N FRUDAKIS,RAODOH MOHAMATH
            PC   C12N15/09,A61K35/14,A61K38/00,A61K39/00,A61K39/39,A61K39/395,
            PC   A61K39/395,
            PC   A61P11/00,A61P35/00,C07K14/47,C07K16/18,C07K19/00,C12N1/19, PC
              C12N1/21,
            PC   C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/574,G01N33/577//
            PC   (C12N1/21,C12R1:19),C12N15/00,A61K37/02,C12N5/00 CC
            Compounds for remedy and diagnosis of lung cancer and method  CC
                  for using the
            CC   same
            FH   Key             Location/Qualifiers
            FT   source          1. .547
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             20.5%;  Score 449;  DB 6;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-239;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy         289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy         349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy         409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy         469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy         529 AGATCTTCGTGAAGACCCTG 548
               ||||||||||||||||||||
Db         528 AGATCTTCGTGAAGACCCTG 547


RESULT 46
AR225519
LOCUS       AR225519                547 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 171 from patent US 6444425.
ACCESSION   AR225519
VERSION     AR225519.1  GI:27263465
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.
  TITLE     Compounds for therapy and diagnosis of lung cancer and methods for
            their use
  JOURNAL   Patent: US 6444425-A 171 03-SEP-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             20.5%;  Score 449;  DB 6;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-239;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy         289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy         349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy         409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy         469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy         529 AGATCTTCGTGAAGACCCTG 548
               ||||||||||||||||||||
Db         528 AGATCTTCGTGAAGACCCTG 547


RESULT 47
AR562942
LOCUS       AR562942                547 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 171 from patent US 6759508.
ACCESSION   AR562942
VERSION     AR562942.1  GI:53977771
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6759508-A 171 06-JUL-2004;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
```

```
     source          1. .547
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             20.5%;  Score 449;  DB 6;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-239;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy         289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy         349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy         409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy         469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy         529 AGATCTTCGTGAAGACCCTG 548
               |||||||||||||||||||
Db         528 AGATCTTCGTGAAGACCCTG 547


RESULT 48
AX321640
LOCUS       AX321640                547 bp    DNA     linear   PAT 15-DEC-2001
DEFINITION  Sequence 171 from Patent WO0172295.
ACCESSION   AX321640
VERSION     AX321640.1  GI:17905852
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
            Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
            Mannion,J. and Kalos,M.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0172295-A 171 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             20.5%;  Score 449;  DB 6;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-239;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy         289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy         349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy         409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy         469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy         529 AGATCTTCGTGAAGACCCTG 548
               |||||||||||||||||||
Db         528 AGATCTTCGTGAAGACCCTG 547


RESULT 49
BD265164
LOCUS       BD265164                460 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compounds for immunotherapy and diagnosis of colonic cancer and
            method of using the same.
ACCESSION   BD265164
VERSION     BD265164.1  GI:33074932
KEYWORDS    JP 2002533082-A/162.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.,
            Wang,T. and Yuqiu,J.
  TITLE     Compounds for immunotherapy and diagnosis of colonic cancer and
            method of using the same
  JOURNAL   Patent: JP 2002533082-A 162 08-OCT-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002533082-A/162
            PD   08-OCT-2002
            PF   23-DEC-1999 JP 2000589697
            PR   23-DEC-1998 US    09/221298,02-JUL-1999 US    09/347496 PR
            22-SEP-1999 US    09/401064,19-NOV-1999 US    09/444242 PR
            02-DEC-1999 US    09/454150
            PI   JIANGCHUN XU,MICHAEL J LODES,HEATHER SECRIST,DARIN R BENSON,
            PI   MADELEINE JOY MEAGHER,JOHN STOLK,TONGTONG WANG,JIANG YUQIU PC
             C12N15/09,A61K31/711,A61K35/14,A61K38/00,A61K39/00,A61K39/395, PC
             A61K39/395,
            PC   A61P35/00,C07K14/47,C07K16/18,C07K19/00,C12N1/15,C12N1/19, PC
```

```
            C12N1/21,
            PC
            C12N5/06,C12N5/10,C12Q1/02,C12Q1/68,G01N33/53,G01N33/53,G01N33/ PC
            566,
            PC   G01N33/574,G01N33/577,G01N33/58,C12N15/00,C12N5/00,C12N5/00,
            PC   A61K37/02
            CC   Compounds for immunotherapy and diagnosis of colonic cancer
            CC          and method of
            CC   using the same
            FH   Key             Location/Qualifiers
            FT   source          1. .460
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             19.8%;  Score 433;  DB 6;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-230;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 50
AR401150
LOCUS       AR401150                460 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 163 from patent US 6623923.
ACCESSION   AR401150
VERSION     AR401150.1  GI:40148450
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Xu,J., Lodes,M.J., Secrist,H., Meagher,M.J., Stolk,J., Benson,D.R.
            and Wang,T.
  TITLE     Compounds for immunotherapy and diagnosis of colon cancer and
            methods for their use
  JOURNAL   Patent: US 6623923-A 163 23-SEP-2003;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             19.8%;  Score 433;  DB 6;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-230;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 51
AX192596
LOCUS       AX192596                460 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 163 from Patent WO0149716.
ACCESSION   AX192596
VERSION     AX192596.1  GI:15210559
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
            Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
  TITLE     Compounds for immunotherapy and diagnosis of colon cancer and
            methods for their use
  JOURNAL   Patent: WO 0149716-A 163 12-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             19.8%;  Score 433;  DB 6;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-230;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460
```

```
RESULT 52
AR225588
LOCUS       AR225588                494 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 275 from patent US 6444425.
ACCESSION   AR225588
VERSION     AR225588.1  GI:27263534
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.
  TITLE     Compounds for therapy and diagnosis of lung cancer and methods for
            their use
  JOURNAL   Patent: US 6444425-A 275 03-SEP-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             19.6%;  Score 430;  DB 6;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 5.1e-229;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494
```

```
RESULT 53
AR563011
LOCUS       AR563011                494 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 275 from patent US 6759508.
ACCESSION   AR563011
VERSION     AR563011.1  GI:53977840
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6759508-A 275 06-JUL-2004;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             19.6%;  Score 430;  DB 6;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 5.1e-229;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373
```

```
Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494


RESULT 54
AX321744
LOCUS       AX321744                494 bp    DNA     linear   PAT 15-DEC-2001
DEFINITION  Sequence 275 from Patent WO0172295.
ACCESSION   AX321744
VERSION     AX321744.1  GI:17906064
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
            Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
            Mannion,J. and Kalos,M.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0172295-A 275 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           19.6%;  Score 430;  DB 6;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 5.1e-229;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494


RESULT 55
BD237089
LOCUS       BD237089                483 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compounds for remedy and diagnosis of lung cancer and method for
            using the same.
ACCESSION   BD237089
VERSION     BD237089.1  GI:33046859
KEYWORDS    JP 2002516659-A/90.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Reed,S.G., Lodes,M.J., Frudakis,T.N. and Mohamath,R.
  TITLE     Compounds for remedy and diagnosis of lung cancer and method for
            using the same
  JOURNAL   Patent: JP 2002516659-A 90 11-JUN-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002516659-A/90
            PD   11-JUN-2002
            PF   26-JAN-1999 JP 2000529432
            PR   28-JAN-1998 US    09/015029,28-JAN-1998 US    09/015022 PR
            18-MAR-1998 US    09/040828,18-MAR-1998 US    09/040831 PR
            23-JUL-1998 US    09/122192,23-JUL-1998 US    09/122191 PR
            22-DEC-1998 US    09/219245
            PI   STEVEN G REED,MICHAEL J LODES,TONY N FRUDAKIS,RAODOH MOHAMATH
            PC   C12N15/09,A61K35/14,A61K38/00,A61K39/00,A61K39/39,A61K39/395,
            PC   A61K39/395,
            PC   A61P11/00,A61P35/00,C07K14/47,C07K16/18,C07K19/00,C12N1/19, PC
              C12N1/21,
            PC   C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/574,G01N33/577//
            PC   (C12N1/21,C12R1:19),C12N15/00,A61K37/02,C12N5/00 CC
            Compounds for remedy and diagnosis of lung cancer and method  CC
                 for using the
            CC   same
            FH   Key             Location/Qualifiers
            FT   source          1. .483
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           19.4%;  Score 426;  DB 6;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 9e-227;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
                ||||||
Db        478 GGTTGA 483


RESULT 56
AR225489
LOCUS       AR225489                483 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 141 from patent US 6444425.
ACCESSION   AR225489
VERSION     AR225489.1  GI:27263435
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.
  TITLE     Compounds for therapy and diagnosis of lung cancer and methods for
            their use
  JOURNAL   Patent: US 6444425-A 141 03-SEP-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            19.4%;  Score 426;  DB 6;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 9e-227;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
                ||||||
Db        478 GGTTGA 483


RESULT 57
AR562912
LOCUS       AR562912                483 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 141 from patent US 6759508.
ACCESSION   AR562912
VERSION     AR562912.1  GI:53977741
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6759508-A 141 06-JUL-2004;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            19.4%;  Score 426;  DB 6;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 9e-227;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
                ||||||
Db        478 GGTTGA 483


RESULT 58
AX321610
LOCUS       AX321610                483 bp    DNA     linear   PAT 15-DEC-2001
DEFINITION  Sequence 141 from Patent WO0172295.
ACCESSION   AX321610
VERSION     AX321610.1  GI:17905756
KEYWORDS    .
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
            Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
            Mannion,J. and Kalos,M.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0172295-A 141 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           19.4%;  Score 426;  DB 6;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 9e-227;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 59
AR563410
LOCUS       AR563410                412 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 692 from patent US 6759508.
ACCESSION   AR563410
VERSION     AR563410.1  GI:53978239
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6759508-A 692 06-JUL-2004;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .412
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           18.8%;  Score 411;  DB 6;  Length 412;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-218;
  Matches  411;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         35 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 61

Qy         95 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 121

Qy        155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 181

Qy        215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 241

Qy        275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 301

Qy        335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 361

Qy        395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 412


RESULT 60
AR225561
LOCUS       AR225561                434 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 248 from patent US 6444425.
ACCESSION   AR225561
VERSION     AR225561.1  GI:27263507
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 434)
  AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.
  TITLE     Compounds for therapy and diagnosis of lung cancer and methods for
            their use
  JOURNAL   Patent: US 6444425-A 248 03-SEP-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           18.7%;  Score 410;  DB 6;  Length 434;
  Best Local Similarity  100.0%;  Pred. No. 8.4e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120
```

```
Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 61
AR562984
LOCUS AR562984 434 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 248 from patent US 6759508.
ACCESSION AR562984
VERSION AR562984.1 GI:53977813
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 434)
AUTHORS Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6759508-A 248 06-JUL-2004;
Corixa Corporation; Seattle, WA
FEATURES Location/Qualifiers
source 1. .434
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 18.7%; Score 410; DB 6; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.4e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 62
AX321717
LOCUS AX321717 434 bp DNA linear PAT 15-DEC-2001
DEFINITION Sequence 248 from Patent WO0172295.
ACCESSION AX321717
VERSION AX321717.1 GI:17905983
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R., Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M., Mannion,J. and Kalos,M.D.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: WO 0172295-A 248 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .434
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 18.7%; Score 410; DB 6; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.4e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 63
CQ696720
LOCUS CQ696720 470 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 41646 from Patent WO02070737.
ACCESSION CQ696720
VERSION CQ696720.1 GI:42246454
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 41646 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           18.3%;  Score 402;  DB 6;  Length 470;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-213;
  Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1717 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 60

Qy       1777 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 120

Qy       1837 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 180

Qy       1897 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 1956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 240

Qy       1957 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGA 2016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGA 300

Qy       2017 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 2076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 360

Qy       2077 AGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGG 2118
             |||||||||||||||||||||||||||||||||||||||||||
Db           361 AGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGG 402


RESULT 64
BD264767
LOCUS       BD264767                518 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compositions and methods for the therapy and diagnosis of ovarian
            cancer.
ACCESSION   BD264767
VERSION     BD264767.1  GI:33074535
KEYWORDS    JP 2002532093-A/152.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A. and Frudakis,T.N.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
  JOURNAL   Patent: JP 2002532093-A 152 02-OCT-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002532093-A/152
            PD   02-OCT-2002
            PF   17-DEC-1999 JP 2000588356
            PR   17-DEC-1998 US    09/215681,17-DEC-1998 US    09/216003 PR
            23-JUN-1999 US    09/338933,24-SEP-1999 US    09/404879 PI
            JENNIFER L MITCHAM,GORDON E KING,PAUL A ALGATE,TONY N FRUDAKIS PC
            C12N15/09,A61K31/7115,A61K35/14,A61K35/76,A61K39/00,A61K39/395, PC
             A61K39/395,
            PC   A61K48/00,A61P35/00,A61P37/04,C07K14/82,C07K19/00,C12N1/15, PC
              C12N1/19,
            PC
            C12N1/21,C12N5/06,C12N5/10,C12Q1/68,G01N33/53,G01N33/53,G01N33/ PC
             566,
            PC   G01N33/574,G01N33/577,C12N15/00,C12N5/00,C12N5/00 CC
            Compositions and methods for the therapy and diagnosis of  CC
                ovarian cancer
            FH   Key             Location/Qualifiers
            FT   source          1. .518
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity  99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db            467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 65
AR238170
LOCUS       AR238170                518 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 152 from patent US 6468546.
ACCESSION   AR238170
VERSION     AR238170.1  GI:27283134
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Mitcham,J.L., King,G.E. and Algate,P.A.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6468546-A 152 22-OCT-2002;
            Corixa Corporation; Seattle, WA
```

```
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 66
AR257711
LOCUS       AR257711                518 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 152 from patent US 6488931.
ACCESSION   AR257711
VERSION     AR257711.1  GI:27307786
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6488931-A 152 03-DEC-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 67
AR283757
LOCUS       AR283757                518 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 152 from patent US 6528253.
ACCESSION   AR283757
VERSION     AR283757.1  GI:29720654
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for diagnosis of ovarian cancer
  JOURNAL   Patent: US 6528253-A 152 04-MAR-2003;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346
```

```
Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 68
AR443431
LOCUS       AR443431                 518 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 152 from patent US 6670463.
ACCESSION   AR443431
VERSION     AR443431.1  GI:42671210
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy of ovarian cancer
  JOURNAL   Patent: US 6670463-A 152 30-DEC-2003;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 69
AR478509
LOCUS       AR478509                 518 bp    DNA     linear   PAT 14-MAY-2004
DEFINITION  Sequence 152 from patent US 6699664.
ACCESSION   AR478509
VERSION     AR478509.1  GI:47237161
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S. and Carter,D.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6699664-A 152 02-MAR-2004;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 70
AR640757
LOCUS       AR640757                 518 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 152 from patent US 6858710.
ACCESSION   AR640757
VERSION     AR640757.1  GI:62775575
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Bangur,C.S., Retter,M.W., Fanger,G.R. and Hill,P.
```

```
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6858710-A 152 22-FEB-2005;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||1||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 71
AX366385
LOCUS       AX366385                518 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 152 from Patent WO0206317.
ACCESSION   AX366385
VERSION     AX366385.1  GI:18697811
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
            Albone,E.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: WO 0206317-A 152 24-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-195;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 72
CQ667521
LOCUS       CQ667521                400 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12447 from Patent WO02070737.
ACCESSION   CQ667521
VERSION     CQ667521.1  GI:42153566
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 12447 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           14.9%;  Score 327;  DB 6;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-171;
  Matches  327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1856 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1915
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 133

Qy       1916 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 1975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        134 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 193

Qy       1976 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 2035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 253

Qy       2036 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 2095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 313

Qy       2096 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 2155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 373

Qy       2156 CATTGCACTTTCCTTTCAATAAAGTTG 2182
              ||||||||||||||||||||||||||
Db        374 CATTGCACTTTCCTTTCAATAAAGTTG 400


RESULT 73
CQ688969
LOCUS       CQ688969                401 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 33895 from Patent WO02070737.
ACCESSION   CQ688969
VERSION     CQ688969.1  GI:42222318
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 33895 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             14.9%;  Score 327;  DB 6;  Length 401;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-171;
  Matches  327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1259 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 60

Qy       1319 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 120

Qy       1379 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 1438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 180

Qy       1439 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 240

Qy       1499 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 300

Qy       1559 GAGGTTGATCTTTGCTGGGAAACAGCT 1585
              ||||||||||||||||||||||||||||
Db        301 GAGGTTGATCTTTGCTGGGAAACAGCT 327


RESULT 74
CQ697121
LOCUS       CQ697121                472 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 42047 from Patent WO02070737.
ACCESSION   CQ697121
VERSION     CQ697121.1  GI:42247916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 42047 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             14.7%;  Score 323;  DB 6;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 5e-169;
  Matches  323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1489 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 60

Qy       1549 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 120

Qy       1609 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 180

Qy       1669 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 240

Qy       1729 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 300

Qy       1789 GGTTGATCTTTGCTGGGAAACAG 1811
              |||||||||||||||||||||||
Db        301 GGTTGATCTTTGCTGGGAAACAG 323


RESULT 75
CQ690613
LOCUS       CQ690613                389 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 35539 from Patent WO02070737.
ACCESSION   CQ690613
VERSION     CQ690613.1  GI:42227675
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 35539 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
```

```
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            13.9%;  Score 304;  DB 6;  Length 389;
  Best Local Similarity  99.7%;  Pred. No. 2.2e-158;
  Matches  354;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1765 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 91

Qy       1825 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 1884
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 151

Qy       1885 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGG 1944
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        152 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCATCCTCGAGGTGG 211

Qy       1945 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 2004
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 271

Qy       2005 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 2064
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 331

Qy       2065 ACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 ACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 386


RESULT 76
CQ684318
LOCUS       CQ684318                308 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 29244 from Patent WO02070737.
ACCESSION   CQ684318
VERSION     CQ684318.1  GI:42209082
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 29244 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .308
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            13.8%;  Score 303;  DB 6;  Length 308;
  Best Local Similarity  100.0%;  Pred. No. 8.1e-158;
  Matches  303;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1888 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 64

Qy       1948 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 2007
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 124

Qy       2008 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2067
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 184

Qy       2068 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2127
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 244

Qy       2128 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2187
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 304

Qy       2188 TTC 2190
             |||
Db        305 TTC 307


RESULT 77
AX193093
LOCUS       AX193093                412 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 660 from Patent WO0149716.
ACCESSION   AX193093
VERSION     AX193093.1  GI:15211048
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
            Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
  TITLE     Compounds for immunotherapy and diagnosis of colon cancer and
            methods for their use
  JOURNAL   Patent: WO 0149716-A 660 12-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .412
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            13.8%;  Score 303;  DB 6;  Length 412;
  Best Local Similarity  99.7%;  Pred. No. 8.1e-158;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 78
CQ689249
LOCUS       CQ689249                312 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34175 from Patent WO02070737.
ACCESSION   CQ689249
VERSION     CQ689249.1  GI:42223197
```

```
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 34175 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .312
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             13.6%;  Score 299;  DB 6;  Length 312;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-155;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1890 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 1949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 73

Qy       1950 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAT 2009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAT 133

Qy       2010 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 2069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 193

Qy       2070 AACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 2129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 253

Qy       2130 TCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCAT 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 TCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCAT 312


RESULT 79
CQ676939
LOCUS       CQ676939                400 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 21865 from Patent WO02070737.
ACCESSION   CQ676939
VERSION     CQ676939.1  GI:42183709
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 21865 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             13.6%;  Score 299;  DB 6;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-155;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1784 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         47 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 106

Qy       1844 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 166

Qy       1904 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 1963
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 226

Qy       1964 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 286

Qy       2024 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 2082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 345


RESULT 80
CQ704345
LOCUS       CQ704345                332 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 49271 from Patent WO02070737.
ACCESSION   CQ704345
VERSION     CQ704345.1  GI:42265114
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 49271 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .332
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             12.8%;  Score 281;  DB 6;  Length 332;
  Best Local Similarity   99.7%;  Pred. No. 1.7e-145;
  Matches  331;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1790 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 60

Qy       1850 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 1909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 120

Qy       1910 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 1969
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCGGTGACACCATCGAGAATGT 180

Qy       1970 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 2029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 240

Qy       2030 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 2089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 300

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 2121
             ||||||||||||||||||||||||||||||||
Db        301 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 332


RESULT 81
```

```
CQ686953
LOCUS       CQ686953                361 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 31879 from Patent WO02070737.
ACCESSION   CQ686953
VERSION     CQ686953.1  GI:42216929
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 31879 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .361
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             12.6%;  Score 276;  DB 6;  Length 361;
  Best Local Similarity   99.7%;  Pred. No. 1.1e-142;
  Matches  326;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 92

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         93 ACCATCACCCTCGAGGTGGAACCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 152

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 2045
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 212

Qy       2046 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 2105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 272

Qy       2106 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 2165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 332

Qy       2166 TCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||
Db        333 TCCTTTCAATAAAGTTGTTGCATTCCC 359


RESULT 82
CQ688744
LOCUS       CQ688744                396 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 33670 from Patent WO02070737.
ACCESSION   CQ688744
VERSION     CQ688744.1  GI:42221690
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 33670 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             12.4%;  Score 271;  DB 6;  Length 396;
  Best Local Similarity   99.7%;  Pred. No. 7e-140;
  Matches  321;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 60

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGNACCCTGTCTGACT 120

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 180

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 240

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 300

Qy       1336 TGATCTTTGCTGGGAAACAGCT 1357
             |||||||||||||||||||||||
Db        301 TGATCTTTGCTGGGAAACAGCT 322


RESULT 83
AR203067
LOCUS       AR203067                284 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 52 from patent US 6365352.
ACCESSION   AR203067
VERSION     AR203067.1  GI:21499359
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 284)
  AUTHORS   Yerramilli,S.V., Prashar,Y., Newburger,P., Goguen,J. and
            Weissman,S.M.
  TITLE     Process to study changes in gene expression in granulocytic cells
  JOURNAL   Patent: US 6365352-A 52 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .284
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             12.3%;  Score 270;  DB 6;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-139;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          15 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 74

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 134

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         135 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 194

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         195 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 254
```

```
Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTG 2182
              ||||||||||||||||||||||||||||||
Db        255 TTTCATTGCACTTTCCTTTCAATAAAGTTG 284


RESULT 84
CQ667335
LOCUS       CQ667335                288 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 12261 from Patent WO02070737.
ACCESSION   CQ667335
VERSION     CQ667335.1  GI:42153031
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 12261 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .288
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           12.3%;  Score 270;  DB 6;  Length 288;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-139;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1898 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACAC 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACAC 78

Qy       1958 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAG 2017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAG 138

Qy       2018 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 2077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 198

Qy       2078 GAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTT 2137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTT 258

Qy       2138 TAAGGTTTCAACAAATTTCATTGCACTTTC 2167
              |||||||||||||||||||||||||||||
Db        259 TAAGGTTTCAACAAATTTCATTGCACTTTC 288


RESULT 85
AY732488
LOCUS       AY732488                490 bp    mRNA    linear   PRI 29-OCT-2004
DEFINITION  Homo sapiens ubiquitin C splice variant (UBC) mRNA, complete cds,
            alternatively spliced.
ACCESSION   AY732488
VERSION     AY732488.1  GI:54300701
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   Wang,P.Z., Wang,F., Chang,Q.S. and Wang,X.
  TITLE     Novel splicing variants of some human genes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 490)
  AUTHORS   Wang,P.Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-AUG-2004) Chinese National Human Genome Center
            Beijing (CHGB), 3# 707 Yongchang Road, BDA, Beijing 100176,
            P.R.China
FEATURES             Location/Qualifiers
     source          1. .490
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .490
                     /gene="UBC"
     CDS             19. .480
                     /gene="UBC"
                     /note="alternatively spliced"
                     /codon_start=1
                     /product="ubiquitin C splice variant"
                     /protein_id="AAV33127.1"
                     /db_xref="GI:54300702"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSGYNIQKESTLHLVLRLRGGV"
ORIGIN

  Query Match           12.1%;  Score 266;  DB 8;  Length 490;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-137;
  Matches  266;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         51 TGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACC 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACC 60

Qy        111 CTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 120

Qy        171 GGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGC 230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGC 180

Qy        231 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 240

Qy        291 GGTGGGATGCAAATCTTCGTGAAGAC 316
              |||||||||||||||||||||||||
Db        241 GGTGGGATGCAAATCTTCGTGAAGAC 266


RESULT 86
CQ698215
LOCUS       CQ698215                345 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 43141 from Patent WO02070737.
ACCESSION   CQ698215
VERSION     CQ698215.1  GI:42251879
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 43141 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           11.8%;  Score 259;  DB 6;  Length 345;
```

```
  Best Local Similarity   100.0%;  Pred. No. 3.7e-133;
  Matches  259;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1854 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 65

Qy       1914 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 1973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 125

Qy       1974 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 2033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 185

Qy       2034 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 2093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 245

Qy       2094 CACTTGGTCCTGCGCTTGA 2112
             ||||||||||||||||||||
Db        246 CACTTGGTCCTGCGCTTGA 264


RESULT 87
CQ686946
LOCUS       CQ686946                339 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 31872 from Patent WO02070737.
ACCESSION   CQ686946
VERSION     CQ686946.1  GI:42216912
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 31872 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .339
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.8%;  Score 258;  DB 6;  Length 339;
  Best Local Similarity   99.7%;  Pred. No. 1.4e-132;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1400 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 90

Qy       1460 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 150

Qy       1520 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 210

Qy       1580 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 270

Qy       1640 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1699
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CCTGGTGCTCCGNCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 330

Qy       1700 CATCACTCT 1708
             |||||||||
Db        331 CATCACTCT 339


RESULT 88
CQ705082
LOCUS       CQ705082                345 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 50008 from Patent WO02070737.
ACCESSION   CQ705082
VERSION     CQ705082.1  GI:42265851
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 50008 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.8%;  Score 258;  DB 6;  Length 345;
  Best Local Similarity   99.7%;  Pred. No. 1.4e-132;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1834 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 1893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 96

Qy       1894 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTG 1953
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 TGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTG 156

Qy       1954 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGC 2013
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGC 216

Qy       2014 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 2073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 276

Qy       2074 TCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCC 2133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCC 336

Qy       2134 CTTTTAAGG 2142
             |||||||||
Db        337 CTTTTAAGG 345


RESULT 89
CQ685697
LOCUS       CQ685697                337 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30623 from Patent WO02070737.
ACCESSION   CQ685697
VERSION     CQ685697.1  GI:42213366
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 30623 12-SEP-2002;
```

```
          Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .337
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.6%;  Score 255;  DB 6;  Length 337;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-131;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1332 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 92

Qy       1392 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 1451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 152

Qy       1452 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 212

Qy       1512 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 272

Qy       1572 GCTGGGAAACAGCTG 1586
             |||||||||||||||
Db        273 GCTGGGAAACAGCTG 287


RESULT 90
CQ923979
LOCUS       CQ923979                543 bp    DNA     linear   PAT 23-NOV-2004
DEFINITION  Sequence 5179 from Patent WO2004097052.
ACCESSION   CQ923979
VERSION     CQ923979.1  GI:56213920
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
            Strahs,A., Immerman,F. and Dorner,A.J.
  TITLE     Methods for prognosis and treatment of solid tumors
  JOURNAL   Patent: WO 2004097052-A 5179 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    34. .60
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    262. .288
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
ORIGIN

  Query Match             11.6%;  Score 255;  DB 6;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-131;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1919 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 1978
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 348

Qy       1979 GATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 2038
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 GATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 408

Qy       2039 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTT 2098
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTT 468

Qy       2099 GGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCAT 2158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCAT 528

Qy       2159 TGCACTTTCCTTTCA 2173
             |||||||||||||||
Db        529 TGCACTTTCCTTTCA 543


RESULT 91
CQ924043
LOCUS       CQ924043                565 bp    DNA     linear   PAT 23-NOV-2004
DEFINITION  Sequence 5243 from Patent WO2004097052.
ACCESSION   CQ924043
VERSION     CQ924043.1  GI:56213984
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L., .
            Strahs,A., Immerman,F. and Dorner,A.J.
  TITLE     Methods for prognosis and treatment of solid tumors
  JOURNAL   Patent: WO 2004097052-A 5243 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES             Location/Qualifiers
     source          1. .565
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.6%;  Score 254;  DB 6;  Length 565;
  Best Local Similarity   99.7%;  Pred. No. 2.4e-130;
  Matches  304;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1860 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 320

Qy       1920 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 1979
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GGTAAGACCATTACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 380

Qy       1980 ATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 440

Qy       2040 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTG 2099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTG 500

Qy       2100 GTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATT 2159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATT 560

Qy       2160 GCACT 2164
             |||||
Db        561 GCACT 565


RESULT 92
BD264660
LOCUS       BD264660                585 bp    DNA     linear   PAT 17-JUL-2003
```

```
DEFINITION  Compositions and methods for the therapy and diagnosis of ovarian
            cancer.
ACCESSION   BD264660
VERSION     BD264660.1  GI:33074428
KEYWORDS    JP 2002532093-A/45.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A. and Frudakis,T.N.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
  JOURNAL   Patent: JP 2002532093-A 45 02-OCT-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002532093-A/45
            PD   02-OCT-2002
            PF   17-DEC-1999 JP 2000588356
            PR   17-DEC-1998 US    09/215681,17-DEC-1998 US    09/216003 PR
            23-JUN-1999 US    09/338933,24-SEP-1999 US    09/404879 PI
            JENNIFER L MITCHAM,GORDON E KING,PAUL A ALGATE,TONY N FRUDAKIS PC
            C12N15/09,A61K31/7115,A61K35/14,A61K35/76,A61K39/00,A61K39/395, PC
             A61K39/395,
            PC   A61K48/00,A61P35/00,A61P37/04,C07K14/82,C07K19/00,C12N1/15, PC
              C12N1/19,
            PC
            C12N1/21,C12N5/06,C12N5/10,C12Q1/68,G01N33/53,G01N33/53,G01N33/ PC
             566,
            PC   G01N33/574,G01N33/577,C12N15/00,C12N5/00,C12N5/00 CC
            Compositions and methods for the therapy and diagnosis of  CC
                ovarian cancer
            FH   Key             Location/Qualifiers
            FT   source          1. .585
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 93
AR238063
LOCUS       AR238063                585 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 45 from patent US 6468546.
ACCESSION   AR238063
VERSION     AR238063.1  GI:27283027
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Mitcham,J.L., King,G.E. and Algate,P.A.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6468546-A 45 22-OCT-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 94
AR257604
LOCUS       AR257604                585 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 45 from patent US 6488931.
ACCESSION   AR257604
VERSION     AR257604.1  GI:27307679
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6488931-A 45 03-DEC-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 95
AR283650
LOCUS       AR283650                585 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 45 from patent US 6528253.
ACCESSION   AR283650
VERSION     AR283650.1  GI:29720547
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for diagnosis of ovarian cancer
  JOURNAL   Patent: US 6528253-A 45 04-MAR-2003;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 96
AR443324
LOCUS       AR443324                585 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 45 from patent US 6670463.
ACCESSION   AR443324
VERSION     AR443324.1  GI:42671103
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy of ovarian cancer
  JOURNAL   Patent: US 6670463-A 45 30-DEC-2003;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 97
AR478402
LOCUS       AR478402                585 bp    DNA     linear   PAT 14-MAY-2004
DEFINITION  Sequence 45 from patent US 6699664.
ACCESSION   AR478402
VERSION     AR478402.1  GI:47237054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S. and Carter,D.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6699664-A 45 02-MAR-2004;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
              ||||||||
Db        532 AAGGTTTC 539


RESULT 98
AR640650
LOCUS       AR640650                585 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 45 from patent US 6858710.
ACCESSION   AR640650
VERSION     AR640650.1  GI:62775468
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Bangur,C.S., Retter,M.W., Fanger,G.R. and Hill,P.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6858710-A 45 22-FEB-2005;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
              ||||||||
Db        532 AAGGTTTC 539


RESULT 99
AX366278
LOCUS       AX366278                585 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 45 from Patent WO0206317.
ACCESSION   AX366278
VERSION     AX366278.1  GI:18697704
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
            Albone,E.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: WO 0206317-A 45 24-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-127;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
              ||||||||
Db        532 AAGGTTTC 539


RESULT 100
CQ514028/c
LOCUS       CQ514028                467 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 45895 from Patent WO0160860.
ACCESSION   CQ514028
VERSION     CQ514028.1  GI:41480292
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Endege,W.O. and Monahan,J.E.
  TITLE     Genes differentially expressed in human prostate cancer and their
            use
  JOURNAL   Patent: WO 0160860-A 45895 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.1%;  Score 244;  DB 6;  Length 467;
  Best Local Similarity   99.4%;  Pred. No. 9.6e-125;
  Matches  344;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        343 AGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCA 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCA 408

Qy        403 TTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 462
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        407 TTCCTCTTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 348

Qy        463 TGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 522
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        347 TGTCTGACTACAACATCCAGAAAGAGTTTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 288

Qy        523 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 228

Qy        583 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATC 168

Qy        643 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        167 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 122


RESULT 101
AX187556
LOCUS       AX187556                286 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 3251 from Patent WO0142467.
ACCESSION   AX187556
VERSION     AX187556.1  GI:15139018.
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 3251 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .286
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.0%;  Score 242;  DB 6;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-123;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 81

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 141

Qy        870 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 201

Qy        930 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 261

Qy        990 TT 991
              ||
Db        262 TT 263


RESULT 102
CQ665885
LOCUS       CQ665885                329 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 10811 from Patent WO02070737.
ACCESSION   CQ665885
VERSION     CQ665885.1  GI:42148359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 10811 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.0%;  Score 242;  DB 6;  Length 329;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-123;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1411 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 60

Qy       1471 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 120

Qy       1531 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 1590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 180

Qy       1591 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 1650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 240

Qy       1651 GT 1652
              ||
Db        241 GT 242


RESULT 103
AX186495
LOCUS       AX186495                487 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 2190 from Patent WO0142467.
ACCESSION   AX186495
VERSION     AX186495.1  GI:15137930
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 2190 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             11.0%;  Score 242;  DB 6;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-123;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 282

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 342

Qy        870 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 402

Qy        930 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 462

Qy        990 TT 991
              ||
Db        463 TT 464


RESULT 104
CQ706133
LOCUS       CQ706133                337 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 51059 from Patent WO02070737.
ACCESSION   CQ706133
VERSION     CQ706133.1  GI:42266902
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 51059 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .337
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            10.9%;  Score 238;  DB 6;  Length 337;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-121;
  Matches  238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 136

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 196

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 2134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 256

Qy       2135 TTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 314


RESULT 105
CQ683715
LOCUS       CQ683715                474 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 28641 from Patent WO02070737.
ACCESSION   CQ683715
VERSION     CQ683715.1  GI:42206865
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 28641 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .474
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            10.8%;  Score 236;  DB 6;  Length 474;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-120;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 95

Qy       1575 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 155

Qy       1635 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 215

Qy       1695 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 271


RESULT 106
CQ688311
LOCUS       CQ688311                286 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 33237 from Patent WO02070737.
ACCESSION   CQ688311
VERSION     CQ688311.1  GI:42411997
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 33237 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .286
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            10.7%;  Score 235;  DB 6;  Length 286;
  Best Local Similarity  99.7%;  Pred. No. 1.1e-119;
  Matches  285;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1337 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 60

Qy       1397 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 GGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 120

Qy       1457 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 180

Qy       1517 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 240

Qy       1577 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
              |||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 286


RESULT 107
CQ682439
LOCUS       CQ682439                237 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 27365 from Patent WO02070737.
ACCESSION   CQ682439
VERSION     CQ682439.1  GI:42202965
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 27365 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .237
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             10.6%;  Score 233;  DB 6;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-118;
  Matches  233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1960 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 2019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 64

Qy       2020 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 2079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 124

Qy       2080 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 2139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 184

Qy       2140 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 237


RESULT 108
CQ689616
LOCUS       CQ689616                232 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34542 from Patent WO02070737.
ACCESSION   CQ689616
VERSION     CQ689616.1  GI:42224349
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 34542 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .232
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             10.6%;  Score 232;  DB 6;  Length 232;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-118;
  Matches  232;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 232


RESULT 109
CQ704983
LOCUS       CQ704983                232 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 49909 from Patent WO02070737.
ACCESSION   CQ704983
VERSION     CQ704983.1  GI:42265752
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 49909 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .232
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             10.2%;  Score 223;  DB 6;  Length 232;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-113;
  Matches  223;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1970 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 2029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 60

Qy       2030 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 2089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 120

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 2149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 180

Qy       2150 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
                ||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 223


RESULT 110
CQ697815
LOCUS       CQ697815                403 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 42741 from Patent WO02070737.
ACCESSION   CQ697815
VERSION     CQ697815.1  GI:42250127
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 42741 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .403
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.9%;  Score 218;  DB 6;  Length 403;
  Best Local Similarity   99.2%;  Pred. No. 3.6e-110;
  Matches  368;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1560 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1619
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 92

Qy       1620 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 1679
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         93 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 152

Qy       1680 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1739
                |||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        153 AAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATCGAGAAT 212

Qy       1740 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1799
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 272

Qy       1800 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1859
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 332

Qy       1860 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1919
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 392

Qy       1920 GGTAAGACCAT 1930
                ||||||||||||
Db        393 GGTAAGACCAT 403


RESULT 111
CQ680957
LOCUS       CQ680957                372 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 25883 from Patent WO02070737.
ACCESSION   CQ680957
VERSION     CQ680957.1  GI:42197453
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 25883 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .372
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.9%;  Score 217;  DB 6;  Length 372;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-109;
  Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 215

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 275

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 335

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
                |||||||||||||||||||||||||||||||||||||
Db        336 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 372


RESULT 112
CQ682331
LOCUS       CQ682331                249 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 27257 from Patent WO02070737.
ACCESSION   CQ682331
VERSION     CQ682331.1  GI:42202577
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 27257 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .249
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.7%;  Score 213;  DB 6;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-107;
  Matches  213;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1425 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 1484
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          37 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 96

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          97 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 156

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         157 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 216
```

```
Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
              ||||||||||||||||||||||||||||||||
Db        217 TCTGACTACAACATCCAGAAAGAGTCCACCCTG 249


RESULT 113
AR240521
LOCUS       AR240521                440 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 277 from patent US 6468758.
ACCESSION   AR240521
VERSION     AR240521.1  GI:27285601
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Benson,D.R., Lodes,M.J., Mitcham,J.L. and King,G.E.
  TITLE     Compositions and methods for ovarian cancer therapy and diagnosis
  JOURNAL   Patent: US 6468758-A 277 22-OCT-2002;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             9.6%;  Score 211;  DB 6;  Length 440;
  Best Local Similarity   100.0%;  Pred. No. 3e-106;
  Matches  211;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 60

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 966
              ||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 211


RESULT 114
CQ686061
LOCUS       CQ686061                261 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30987 from Patent WO02070737.
ACCESSION   CQ686061
VERSION     CQ686061.1  GI:42214378
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 30987 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .261
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.6%;  Score 210;  DB 6;  Length 261;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-105;
  Matches  260;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1130 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 60

Qy       1190 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 120

Qy       1250 TCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1309
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 TCTCGAGGTGGAGCCGAGTGACCCCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 180

Qy       1310 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 240

Qy       1370 CACCCTGTCTGACTACAACAT 1390
              |||||||||||||||||||||
Db        241 CACCCTGTCTGACTACAACAT 261


RESULT 115
CQ663336
LOCUS       CQ663336                330 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8262 from Patent WO02070737.
ACCESSION   CQ663336
VERSION     CQ663336.1  GI:42141591
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 8262 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .330
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.6%;  Score 210;  DB 6;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-105;
  Matches  210;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 101

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 161

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 221

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGG 226
              ||||||||||||||||||||||||||||||
Db        222 GATCTTTGCTGGAAAACAGCTGGAAGATGG 251


RESULT 116
AR606553
LOCUS       AR606553                228 bp    DNA     linear   PAT 15-DEC-2004
DEFINITION  Sequence 80 from patent US 6821519.
```

```
ACCESSION   AR606553
VERSION     AR606553.1  GI:56658515
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 228)
  AUTHORS   Day,C.H., Hosken,N.A. and Parsons,J.M.
  TITLE     Compositions and methods for the diagnosis and treatment of herpes
            simplex virus infection
  JOURNAL   Patent: US 6821519-A 80 23-NOV-2004;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .228
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             9.4%;  Score 207;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-104;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              ||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 117
BD229494/c
LOCUS       BD229494                609 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human gene differentially expressed in colonic cancer.
ACCESSION   BD229494
VERSION     BD229494.1  GI:33039264
KEYWORDS    JP 2002523088-A/169.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 609)
  AUTHORS   Endege,W.O., Steinmann,K.E., Astle,J.H., Burgess,C.C., Iii,E.C.,
            Catino,T.J., Dwivedi,P., Ford,D.M., Lewis,M.E., Molino,G.A.,
            Monahan,J.E. and Schlegel,R.
  TITLE     Human gene differentially expressed in colonic cancer
  JOURNAL   Patent: JP 2002523088-A 169 30-JUL-2002;
            BAYER CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002523088-A/169
            PD   30-JUL-2002
            PF   30-AUG-1999 JP 2000567689
            PR   31-AUG-1998 US    60/098639,27-JAN-1999 US    60/117393 PI
            WILSON O ENDEGE,KATHLEEN E STEINMANN,JOHN H ASTLE,CHRISTOPHER  PI
            C BURGESS,
            PI   EDDIE CARROLL III,THEODORE J CATINO,POORNIMA DWIVEDI,DONNA M
            PI   FORD,
            PI   MARCIA E LEWIS,GARY A MOLINO,JOHN E MONAHAN,ROBERT SCHLEGEL PC
              C12N15/09,C12N15/09,A01K67/027,A61K38/00,A61K45/00,A61K48/00, PC
             A61P35/00,
            PC   C07K14/82,C07K16/32,C12M1/00,C12M1/40,C12N1/15,C12N1/19,C12N1/
            PC   21,C12N5/10,
            PC   C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/574//C12P21/08,
            PC   C12N15/00,
            PC   C12N5/00,C12N15/00,A61K37/02
            CC   n = A,T,C or G
            FH   Key             Location/Qualifiers
            FT   misc_feature    (1). .(609).
FEATURES             Location/Qualifiers
     source          1. .609
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.4%;  Score 207;  DB 6;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-104;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 305

Qy       1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 245

Qy       1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 185

Qy       1871 GGTGCTCCGTCTCAGAGGTGGGATGCA 1897
              ||||||||||||||||||||||||||
Db        184 GGTGCTCCGTCTCAGAGGTGGGATGCA 158


RESULT 118
CQ658851
LOCUS       CQ658851                384 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 3777 from Patent WO02070737.
ACCESSION   CQ658851
VERSION     CQ658851.1  GI:42125649
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 3777 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .384
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.3%;  Score 204;  DB 6;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-102;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1700 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 60

Qy       1760 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 120

Qy       1820 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 180
```

```
Qy       1880 TCTCAGAGGTGGGATGCAGATCTT 1903
             ||||||||||||||||||||||||
Db        181 TCTCAGAGGTGGGATGCAGATCTT 204


RESULT 119
AX816038
LOCUS       AX816038                480 bp    DNA     linear   PAT 09-DEC-2003
DEFINITION  Sequence 9 from Patent WO03066660.
ACCESSION   AX816038
VERSION     AX816038.1  GI:39646658
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Caterer,N.R., Uttental,L.O. and Neilson,R.W.
  TITLE     Compositions and methods for assembly and stabilization of antibody
            fv fragments via antiparallel heterogeneous coiled&shy;coil peptide
            regions and uses thereof
  JOURNAL   Patent: WO 03066660-A 9 14-AUG-2003;
            Immunolex Therapeutics APS (DK)
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

  Query Match             9.3%;  Score 203;  DB 6;  Length 480;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-102;
  Matches  203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 156

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 216

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 276

Qy        726 CACCTGGTACTCCGTCTCAGAGG 748
              |||||||||||||||||||||||
Db        277 CACCTGGTACTCCGTCTCAGAGG 299


RESULT 120
AX816040
LOCUS       AX816040                504 bp    DNA     linear   PAT 09-DEC-2003
DEFINITION  Sequence 11 from Patent WO03066660.
ACCESSION   AX816040
VERSION     AX816040.1  GI:39646659
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Caterer,N.R., Uttental,L.O. and Neilson,R.W.
  TITLE     Compositions and methods for assembly and stabilization of antibody
            fv fragments via antiparallel heterogeneous coiled&shy;coil peptide
            regions and uses thereof
  JOURNAL   Patent: WO 03066660-A 11 14-AUG-2003;
            Immunolex Therapeutics APS (DK)
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
ORIGIN

  Query Match             9.3%;  Score 203;  DB 6;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-102;
  Matches  203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 180

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 240

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 300

Qy        726 CACCTGGTACTCCGTCTCAGAGG 748
              |||||||||||||||||||||||
Db        301 CACCTGGTACTCCGTCTCAGAGG 323


RESULT 121
CQ691540
LOCUS       CQ691540                261 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 36466 from Patent WO02070737.
ACCESSION   CQ691540
VERSION     CQ691540.1  GI:42230685
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 36466 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .261
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.2%;  Score 202;  DB 6;  Length 261;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-101;
  Matches  202;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1989 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 117

Qy       2049 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 177

Qy       2109 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCC 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCC 237

Qy       2169 TTTCAATAAAGTTGTTGCATTC 2190
              |||||||||||||||||||||||
Db        238 TTTCAATAAAGTTGTTGCATTC 259


RESULT 122
CQ695149
LOCUS       CQ695149                490 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 40075 from Patent WO02070737.
```

```
ACCESSION   CQ695149
VERSION     CQ695149.1  GI:42241029
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 40075 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .490
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.2%;  Score 202;  DB 6;  Length 490;
  Best Local Similarity   99.3%;  Pred. No. 3.4e-101;
  Matches  302;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1787 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 1846
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 169

Qy       1847 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGT 1906
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        170 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGNGGGATGCAGATCTTCGT 229

Qy       1907 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 1966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 289

Qy       1967 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 2026
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        290 TGTCAAGGCAAAGATCCAAGATAAGGAGGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 349

Qy       2027 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 2086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 409

Qy       2087 CACT 2090
             ||||
Db        410 CACT 413


RESULT 123
CQ923956
LOCUS       CQ923956                549 bp    DNA     linear   PAT 23-NOV-2004
DEFINITION  Sequence 5156 from Patent WO2004097052.
ACCESSION   CQ923956
VERSION     CQ923956.1  GI:56213897
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
            Strahs,A., Immerman,F. and Dorner,A.J.
  TITLE     Methods for prognosis and treatment of solid tumors
  JOURNAL   Patent: WO 2004097052-A 5156 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    108. .134
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    336. .362
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
ORIGIN

  Query Match             9.2%;  Score 201;  DB 6;  Length 549;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 194

Qy       1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 254

Qy       1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 314

Qy       1871 GGTGCTCCGTCTCAGAGGTGG 1891
             |||||||||||||||||||||
Db        315 GGTGCTCCGTCTCAGAGGTGG 335


RESULT 124
CQ690997
LOCUS       CQ690997                216 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 35923 from Patent WO02070737.
ACCESSION   CQ690997
VERSION     CQ690997.1  GI:42229029
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 35923 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .216
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.1%;  Score 200;  DB 6;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-100;
  Matches  200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2052
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          17 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 76

Qy       2053 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          77 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 136

Qy       2113 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2172
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         137 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 196

Qy       2173 AATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||
Db         197 AATAAAGTTGTTGCATTCCC 216
```

```
RESULT 125
CQ668894
LOCUS       CQ668894                367 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 13820 from Patent WO02070737.
ACCESSION   CQ668894
VERSION     CQ668894.1  GI:42157634
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 13820 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .367
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             9.1%;  Score 200;  DB 6;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-100;
  Matches  200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 154

Qy       1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 214

Qy       1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 274

Qy       1689 ACTGGTAAGACCATCACTCT 1708
             ||||||||||||||||||||
Db        275 ACTGGTAAGACCATCACTCT 294


RESULT 126
CQ705824
LOCUS       CQ705824                294 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 50750 from Patent WO02070737.
ACCESSION   CQ705824
VERSION     CQ705824.1  GI:42266593
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 50750 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.8%;  Score 192;  DB 6;  Length 294;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-95;
  Matches  292;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 60

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCNGATCAGCAGAGG 120

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 TTGATCTTTGCTGGGAAACAGCTGNAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 240

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 294


RESULT 127
BD264710/c
LOCUS       BD264710                605 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compositions and methods for the therapy and diagnosis of ovarian
            cancer.
ACCESSION   BD264710
VERSION     BD264710.1  GI:33074478
KEYWORDS    JP 2002532093-A/95.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A. and Frudakis,T.N.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
  JOURNAL   Patent: JP 2002532093-A 95 02-OCT-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002532093-A/95
            PD   02-OCT-2002
            PF   17-DEC-1999 JP 2000588356
            PR   17-DEC-1998 US    09/215681,17-DEC-1998 US    09/216003 PR
            23-JUN-1999 US    09/338933,24-SEP-1999 US    09/404879 PI
            JENNIFER L MITCHAM,GORDON E KING,PAUL A ALGATE,TONY N FRUDAKIS PC
            C12N15/09,A61K31/7115,A61K35/14,A61K35/76,A61K39/00,A61K39/395, PC
             A61K39/395,
            PC   A61K48/00,A61P35/00,A61P37/04,C07K14/82,C07K19/00,C12N1/15, PC
              C12N1/19,
            PC
            C12N1/21,C12N5/06,C12N5/10,C12Q1/68,G01N33/53,G01N33/53,G01N33/ PC
             566,
            PC   G01N33/574,G01N33/577,C12N15/00,C12N5/00,C12N5/00 CC
            Compositions and methods for the therapy and diagnosis of  CC
                ovarian cancer
            FH   Key             Location/Qualifiers
            FT   source          1. .605
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300
```

```
Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 128
AR238113/c
LOCUS       AR238113                605 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 95 from patent US 6468546.
ACCESSION   AR238113
VERSION     AR238113.1  GI:27283077
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mitcham,J.L., King,G.E. and Algate,P.A.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6468546-A 95 22-OCT-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 129
AR257654/c
LOCUS       AR257654                605 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 95 from patent US 6488931.
ACCESSION   AR257654
VERSION     AR257654.1  GI:27307729
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6488931-A 95 03-DEC-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 130
AR283700/c
LOCUS       AR283700                605 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 95 from patent US 6528253.
ACCESSION   AR283700
VERSION     AR283700.1  GI:29720597
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for diagnosis of ovarian cancer
  JOURNAL   Patent: US 6528253-A 95 04-MAR-2003;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 131
AR443374/c
LOCUS       AR443374                605 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 95 from patent US 6670463.
```

```
ACCESSION   AR443374
VERSION     AR443374.1  GI:42671153
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy of ovarian cancer
  JOURNAL   Patent: US 6670463-A 95 30-DEC-2003;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 132
AR478452/c
LOCUS       AR478452                605 bp    DNA     linear   PAT 14-MAY-2004
DEFINITION  Sequence 95 from patent US 6699664.
ACCESSION   AR478452
VERSION     AR478452.1  GI:47237104
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S. and Carter,D.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6699664-A 95 02-MAR-2004;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 133
AR640700/c
LOCUS       AR640700                605 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 95 from patent US 6858710.
ACCESSION   AR640700
VERSION     AR640700.1  GI:62775518
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Bangur,C.S., Retter,M.W., Fanger,G.R. and Hill,P.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6858710-A 95 22-FEB-2005;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 134
AX366328/c
LOCUS       AX366328                605 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 95 from Patent WO0206317.
ACCESSION   AX366328
VERSION     AX366328.1  GI:18697754
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
            Albone,E.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
```

```
  JOURNAL   Patent: WO 0206317-A 95 24-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 135
BD264715
LOCUS       BD264715                 697 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compositions and methods for the therapy and diagnosis of ovarian
            cancer.
ACCESSION   BD264715
VERSION     BD264715.1  GI:33074483
KEYWORDS    JP 2002532093-A/100.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A. and Frudakis,T.N.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
  JOURNAL   Patent: JP 2002532093-A 100 02-OCT-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002532093-A/100
            PD   02-OCT-2002
            PF   17-DEC-1999 JP 2000588356
            PR   17-DEC-1998 US    09/215681,17-DEC-1998 US    09/216003 PR
            23-JUN-1999 US    09/338933,24-SEP-1999 US    09/404879 PI
            JENNIFER L MITCHAM,GORDON E KING,PAUL A ALGATE,TONY N FRUDAKIS PC
            C12N15/09,A61K31/7115,A61K35/14,A61K35/76,A61K39/00,A61K39/395, PC
             A61K39/395,
            PC   A61K48/00,A61P35/00,A61P37/04,C07K14/82,C07K19/00,C12N1/15, PC
              C12N1/19,
            PC
            C12N1/21,C12N5/06,C12N5/10,C12Q1/68,G01N33/53,G01N33/53,G01N33/ PC
             566,
            PC   G01N33/574,G01N33/577,C12N15/00,C12N5/00,C12N5/00 CC
            Compositions and methods for the therapy and diagnosis of  CC
                ovarian cancer
            FH   Key             Location/Qualifiers
            FT   source          1. .697
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 136
AR238118
LOCUS       AR238118                 697 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 100 from patent US 6468546.
ACCESSION   AR238118
VERSION     AR238118.1  GI:27283082
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Mitcham,J.L., King,G.E. and Algate,P.A.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6468546-A 100 22-OCT-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 137
AR257659
LOCUS       AR257659                 697 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 100 from patent US 6488931.
ACCESSION   AR257659
VERSION     AR257659.1  GI:27307734
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6488931-A 100 03-DEC-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 138
AR283705
LOCUS       AR283705                697 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 100 from patent US 6528253.
ACCESSION   AR283705
VERSION     AR283705.1  GI:29720602
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for diagnosis of ovarian cancer
  JOURNAL   Patent: US 6528253-A 100 04-MAR-2003;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 139
AR443379
LOCUS       AR443379                697 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 100 from patent US 6670463.
ACCESSION   AR443379
VERSION     AR443379.1  GI:42671158
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy of ovarian cancer
  JOURNAL   Patent: US 6670463-A 100 30-DEC-2003;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 140
AR478457
LOCUS       AR478457                697 bp    DNA     linear   PAT 14-MAY-2004
DEFINITION  Sequence 100 from patent US 6699664.
ACCESSION   AR478457
VERSION     AR478457.1  GI:47237109
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S. and Carter,D.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6699664-A 100 02-MAR-2004;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
```

```
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
              ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 141
AR640705
LOCUS       AR640705                 697 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 100 from patent US 6858710.
ACCESSION   AR640705
VERSION     AR640705.1  GI:62775523
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Bangur,C.S., Retter,M.W., Fanger,G.R. and Hill,P.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6858710-A 100 22-FEB-2005;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
              ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 142
AX366333
LOCUS       AX366333                 697 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 100 from Patent WO0206317.
ACCESSION   AX366333
VERSION     AX366333.1  GI:18697759
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
            Albone,E.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: WO 0206317-A 100 24-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-95;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
              ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 143
AX207733
LOCUS       AX207733                 6340 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 31 from Patent WO0157242.
ACCESSION   AX207733
VERSION     AX207733.1  GI:15422417
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 31 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6340
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-GFP construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6340;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044
```

```
Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 144
AX207725
LOCUS       AX207725               6411 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 23 from Patent WO0157242.
ACCESSION   AX207725
VERSION     AX207725.1  GI:15422409
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 23 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6411
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6411;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 145
AX207729
LOCUS       AX207729               6411 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 27 from Patent WO0157242.
ACCESSION   AX207729
VERSION     AX207729.1  GI:15422413
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 27 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6411
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-Ub-Met-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6411;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 146
AX207740
LOCUS       AX207740               6436 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 38 from Patent WO0157242.
ACCESSION   AX207740
VERSION     AX207740.1  GI:15422424
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 38 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6436
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-C3 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6436;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 147
AX207749
LOCUS       AX207749               6459 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 47 from Patent WO0157242.
```

```
ACCESSION   AX207749
VERSION     AX207749.1  GI:15422433
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 47 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6459
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-DEVD-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6459;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 148
AX207755
LOCUS       AX207755                6459 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 53 from Patent WO0157242.
ACCESSION   AX207755
VERSION     AX207755.1  GI:15422439
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 53 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6459
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-1XUb-DEVA-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6459;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 149
AX207734
LOCUS       AX207734                6607 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 32 from Patent WO0157242.
ACCESSION   AX207734
VERSION     AX207734.1  GI:15422418
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 32 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6607
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-GFP construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6607;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 150
AX207726
LOCUS       AX207726                6678 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 24 from Patent WO0157242.
ACCESSION   AX207726
VERSION     AX207726.1  GI:15422410
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 24 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6678
                     /organism="synthetic construct"
```

```
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6678;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 151
AX207741
LOCUS       AX207741               6703 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 39 from Patent WO0157242.
ACCESSION   AX207741
VERSION     AX207741.1  GI:15422425
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 39 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6703
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-C3 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6703;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 152
AX207750
LOCUS       AX207750               6726 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 48 from Patent WO0157242.
ACCESSION   AX207750
VERSION     AX207750.1  GI:15422434
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 48 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6726
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-DEVD-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6726;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 153
AX207756
LOCUS       AX207756               6726 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 54 from Patent WO0157242.
ACCESSION   AX207756
VERSION     AX207756.1  GI:15422440
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 54 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6726
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-2XUb-DEVA-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6726;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
                ||||||||
Db       1129 CACCTGGT 1136


RESULT 154
AX207735
LOCUS       AX207735                6850 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 33 from Patent WO0157242.
ACCESSION   AX207735
VERSION     AX207735.1  GI:15422419
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 33 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6850
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-GFP construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6850;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
                ||||||||
Db       1129 CACCTGGT 1136


RESULT 155
AX207727
LOCUS       AX207727                6921 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 25 from Patent WO0157242.
ACCESSION   AX207727
VERSION     AX207727.1  GI:15422411
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 25 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6921
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6921;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
                ||||||||
Db       1129 CACCTGGT 1136


RESULT 156
AX207742
LOCUS       AX207742                6946 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 40 from Patent WO0157242.
ACCESSION   AX207742
VERSION     AX207742.1  GI:15422426
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 40 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6946
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-C3 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6946;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
                ||||||||
Db       1129 CACCTGGT 1136


RESULT 157
AX207751
```

```
LOCUS       AX207751               6969 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 49 from Patent WO0157242.
ACCESSION   AX207751
VERSION     AX207751.1  GI:15422435
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 49 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6969
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-DEVD-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6969;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 158
AX207757
LOCUS       AX207757               6969 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 55 from Patent WO0157242.
ACCESSION   AX207757
VERSION     AX207757.1  GI:15422441
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 55 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .6969
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-DEVA-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 6969;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 159
AX207767
LOCUS       AX207767               7053 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 65 from Patent WO0157242.
ACCESSION   AX207767
VERSION     AX207767.1  GI:15422451
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 65 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7053
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-Ub-Met-Bla HRV16 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7053;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 160
AX207772
LOCUS       AX207772               7053 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 70 from Patent WO0157242.
ACCESSION   AX207772
VERSION     AX207772.1  GI:15422456
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 70 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
```

```
     source          1. .7053
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-Ub-Met-Bla HRV16 (C106A) construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7053;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 161
AX207773
LOCUS       AX207773                7053 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 71 from Patent WO0157242.
ACCESSION   AX207773
VERSION     AX207773.1  GI:15422457
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 71 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7053
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-Ub-Met-Bla HRV16 (D35A) construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7053;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 162
AX207736
LOCUS       AX207736                7093 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 34 from Patent WO0157242.
ACCESSION   AX207736
VERSION     AX207736.1  GI:15422420
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 34 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7093
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-GFP construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7093;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 163
AX207728
LOCUS       AX207728                7164 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 26 from Patent WO0157242.
ACCESSION   AX207728
VERSION     AX207728.1  GI:15422412
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 26 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7164
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7164;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008
```

```
Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 164
AX207743
LOCUS       AX207743                7189 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 41 from Patent WO0157242.
ACCESSION   AX207743
VERSION     AX207743.1  GI:15422427
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 41 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7189
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-C3 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7189;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 165
AX207752
LOCUS       AX207752                7212 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 50 from Patent WO0157242.
ACCESSION   AX207752
VERSION     AX207752.1  GI:15422436
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 50 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7212
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-DEVD-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7212;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 166
AX207758
LOCUS       AX207758                7212 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 56 from Patent WO0157242.
ACCESSION   AX207758
VERSION     AX207758.1  GI:15422442
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 56 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7212
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-4XUb-DEVA-Bla construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7212;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136
```

```
RESULT 167
AX207746
LOCUS       AX207746               7248 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 44 from Patent WO0157242.
ACCESSION   AX207746
VERSION     AX207746.1  GI:15422430
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 44 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7248
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-Ub-Met-C3 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7248;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 168
AX207774
LOCUS       AX207774               7512 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 72 from Patent WO0157242.
ACCESSION   AX207774
VERSION     AX207774.1  GI:15422458
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 72 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7512
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-MetUb-Bla HR14 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7512;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 169
AX207766
LOCUS       AX207766               7563 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 64 from Patent WO0157242.
ACCESSION   AX207766
VERSION     AX207766.1  GI:15422450
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 64 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7563
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-Bla HRV16 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7563;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 170
AX207770
LOCUS       AX207770               7563 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 68 from Patent WO0157242.
ACCESSION   AX207770
VERSION     AX207770.1  GI:15422454
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 68 09-AUG-2001;
```

```
          Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7563
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-Bla HRV16 (C106A) construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7563;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 171
AX207771
LOCUS       AX207771               7563 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 69 from Patent WO0157242.
ACCESSION   AX207771
VERSION     AX207771.1  GI:15422455
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 69 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .7563
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-Bla HRV16 (D35A) construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 7563;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 172
AX207762
LOCUS       AX207762               8022 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 60 from Patent WO0157242.
ACCESSION   AX207762
VERSION     AX207762.1  GI:15422446
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 60 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .8022
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="pcDNA3-3XUb-Bla HRV 14 construct"
ORIGIN

  Query Match             8.6%;  Score 188;  DB 6;  Length 8022;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-93;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 173
CQ679610
LOCUS       CQ679610               355 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 24536 from Patent WO02070737.
ACCESSION   CQ679610
VERSION     CQ679610.1  GI:42192639
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 24536 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.3%;  Score 183;  DB 6;  Length 355;
  Best Local Similarity   99.3%;  Pred. No. 1.5e-90;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 60

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954
             ||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 GCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGNGGAGCCCAGTGA 120

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 180

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 240

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 285


RESULT 174
CQ660774
LOCUS       CQ660774                321 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 5700 from Patent WO02070737.
ACCESSION   CQ660774
VERSION     CQ660774.1  GI:42132722
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 5700 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.3%;  Score 182;  DB 6;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-90;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1853 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 140

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         141 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 200

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 260

Qy       2033 GA 2034
             ||
Db         261 GA 262


RESULT 175
AX209208
LOCUS       AX209208                227 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 1048 from Patent WO0157207.
ACCESSION   AX209208
VERSION     AX209208.1  GI:15423631
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Algate,P.A. and Mannion,J.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: WO 0157207-A 1048 09-AUG-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .227
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.0%;  Score 176;  DB 6;  Length 227;
  Best Local Similarity   99.6%;  Pred. No. 1.3e-86;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 60

Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 TGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 120

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 180

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 227


RESULT 176
AX198675
LOCUS       AX198675                335 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 1130 from Patent WO0151513.
ACCESSION   AX198675
VERSION     AX198675.1  GI:15388996
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Algate,P.A.
  TITLE     Ovarian tumor-associated sequences
  JOURNAL   Patent: WO 0151513-A 1130 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .335
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             8.0%;  Score 176;  DB 6;  Length 335;
  Best Local Similarity   99.6%;  Pred. No. 1.3e-86;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 60

Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
```

```
Db         61 TGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 120

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 180

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 227


RESULT 177
CQ484202/c
LOCUS       CQ484202                294 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 16069 from Patent WO0160860.
ACCESSION   CQ484202
VERSION     CQ484202.1  GI:41449821
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Endege,W.O. and Monahan,J.E.
  TITLE     Genes differentially expressed in human prostate cancer and their
            use
  JOURNAL   Patent: WO 0160860-A 16069 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             7.8%;  Score 171;  DB 6;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 8e-84;
  Matches  171;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        518 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 203

Qy        578 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCC 637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCC 143

Qy        638 TGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 92


RESULT 178
AX207704
LOCUS       AX207704                228 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 2 from Patent WO0157242.
ACCESSION   AX207704
VERSION     AX207704.1  GI:15422383
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 2 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .228
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
                     /note="Eukaryote"
ORIGIN

  Query Match             7.6%;  Score 167;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-81;
  Matches  167;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1479 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 102

Qy       1539 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 162

Qy       1599 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 209


RESULT 179
AX207719
LOCUS       AX207719                228 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 17 from Patent WO0157242.
ACCESSION   AX207719
VERSION     AX207719.1  GI:15422403
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.
  TITLE     Methods of protein destabilization and uses thereof
  JOURNAL   Patent: WO 0157242-A 17 09-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .228
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="UbquitinG76V mutant"
ORIGIN

  Query Match             7.6%;  Score 167;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-81;
  Matches  167;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1479 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 102

Qy       1539 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 162

Qy       1599 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 209


RESULT 180
CQ691600
LOCUS       CQ691600                273 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 36526 from Patent WO02070737.
ACCESSION   CQ691600
VERSION     CQ691600.1  GI:42230880
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 36526 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .273
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             7.3%;  Score 161;  DB 6;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-78;
  Matches  161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         34 TCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGA 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 TCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGA 68

Qy         94 CTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 128

Qy        154 AGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              |||||||||||||||||||||||||||||||||||||||||
Db        129 AGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 169


RESULT 181
CQ705545
LOCUS       CQ705545                 211 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 50471 from Patent WO02070737.
ACCESSION   CQ705545
VERSION     CQ705545.1  GI:42266314
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 50471 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .211
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             7.3%;  Score 160;  DB 6;  Length 211;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-77;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 111

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 171

Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              |||||||||||||||||||||||||||||||||||||||
Db        172 TTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 211


RESULT 182
BV207846
LOCUS       BV207846                 201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm224695 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV207846
VERSION     BV207846.1  GI:48178451
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             6.8%;  Score 150;  DB 10;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 4.7e-72;
  Matches  200;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1702 TCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 1761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 60

Qy       1762 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 1821
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCYGGGAAACAGCTGGAAGATG 120

Qy       1822 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 1881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 180

Qy       1882 TCAGAGGTGGGATGCAGATCT 1902
              |||||||||||||||||||||
Db        181 TCAGAGGTGGGATGCAGATCT 201


RESULT 183
CQ663822
LOCUS       CQ663822                 400 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8748 from Patent WO02070737.
ACCESSION   CQ663822
VERSION     CQ663822.1  GI:42143376
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 8748 12-SEP-2002;
            Chondrogene Inc. (CA)
```

```
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             6.5%;  Score 143;  DB 6;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 4e-68;
  Matches  143;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 176

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 236

Qy       2019 TTGATCTTTGCTGGGAAACAGCT 2041
             |||||||||||||||||||||||
Db        237 TTGATCTTTGCTGGGAAACAGCT 259


RESULT 184
AB089615
LOCUS       AB089615               1830 bp    DNA     linear   PRI 03-FEB-2004
DEFINITION  Gorilla gorilla UbC gene for polyubiquitin C, complete cds.
ACCESSION   AB089615
VERSION     AB089615.1  GI:28316339
KEYWORDS    .
SOURCE      Gorilla gorilla (gorilla)
  ORGANISM  Gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Gorilla.
REFERENCE   1
  AUTHORS   Tachikui,H., Saitou,N., Nakajima,T., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Lineage-specific homogenization of the polyubiquitin gene among
            human and great apes
  JOURNAL   J. Mol. Evol. 57 (6), 737-744 (2003)
   PUBMED   14745543
REFERENCE   2  (bases 1 to 1830)
  AUTHORS   Tachikui,H., Nakajima,T., Saito,N., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2002) Hiroshi Tachikui, Institute of Medical
            Science, University of tokyo, Division of Genetic Dignosis;
            Shirokanedai 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:tachikui@ims.u-tokyo.ac.jp, Tel:81-5449-5750,
            Fax:81-5449-5764)
FEATURES             Location/Qualifiers
     source          1. .1830
                     /organism="Gorilla gorilla"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9593"
                     /note="Gorilla with 8 ub units"
     gene            1. .1830
                     /gene="UbC"
     CDS             1. .1830
                     /gene="UbC"
                     /codon_start=1
                     /product="polyubiquitin C"
                     /protein_id="BAC56953.1"
                     /db_xref="GI:28316340"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGV"
ORIGIN

  Query Match             6.5%;  Score 143;  DB 8;  Length 1830;
  Best Local Similarity   100.0%;  Pred. No. 4e-68;
  Matches  143;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1824 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 1359

Qy       1884 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 1943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 1419

Qy       1944 GAGCCCAGTGACACCATCGAGAA 1966
             |||||||||||||||||||||||
Db       1420 GAGCCCAGTGACACCATCGAGAA 1442


RESULT 185
BV200498
LOCUS       BV200498               201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm204274 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV200498
VERSION     BV200498.1  GI:48169078
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             6.5%;  Score 142;  DB 10;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 1.4e-67;
  Matches  192;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1233 ACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1292
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 ACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 68

Qy       1293 AAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAA 1352
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         69 AAGATCCAAGACAAGGAAGGCATCCCTCCTGAYCAGCAGAGGTTGATCTTTGCTGGGAAA 128
```

```
Qy       1353 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 188

Qy       1413 CTGGTGCTCCGTC 1425
              |||||||||||||
Db        189 CTGGTGCTCCGTC 201


RESULT 186
BV200925
LOCUS       BV200925                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm205429 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV200925
VERSION     BV200925.1  GI:48169610
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             6.5%;  Score 142;  DB 10;  Length 201;
  Best Local Similarity   99.5%;  Pred. No. 1.4e-67;
  Matches  192;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1098 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 68

Qy       1158 AACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1217
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         69 AACATCCAGAAAGAGTCCACCTTGCACCTGGTRCTCCGTCTCAGAGGTGGGATGCAGATC 128

Qy       1218 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATT 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATT 188

Qy       1278 GAGAATGTCAAGG 1290
              |||||||||||||
Db        189 GAGAATGTCAAGG 201


RESULT 187
AX198840/c
LOCUS       AX198840                240 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 1295 from Patent WO0151513.
ACCESSION   AX198840
VERSION     AX198840.1  GI:15389161
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Algate,P.A.
  TITLE     Ovarian tumor-associated sequences
  JOURNAL   Patent: WO 0151513-A 1295 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .240
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             6.2%;  Score 137;  DB 6;  Length 240;
  Best Local Similarity   99.5%;  Pred. No. 9.2e-65;
  Matches  187;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 129

Qy       1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        128 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 69

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 9

Qy       1575 GGGAAACA 1582
              ||||||||
Db           8 GGGAAACA 1


RESULT 188
AX209367/c
LOCUS       AX209367                240 bp    DNA     linear   PAT 31-AUG-2001
DEFINITION  Sequence 1207 from Patent WO0157207.
ACCESSION   AX209367
VERSION     AX209367.1  GI:15423790
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Algate,P.A. and Mannion,J.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: WO 0157207-A 1207 09-AUG-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .240
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             6.2%;  Score 137;  DB 6;  Length 240;
  Best Local Similarity   99.5%;  Pred. No. 9.2e-65;
  Matches  187;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        188 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 129

Qy       1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        128 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 69

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 9

Qy       1575 GGGAAACA 1582
              ||||||||
Db          8 GGGAAACA 1
```

```
RESULT 189
AB089616
LOCUS       AB089616               2286 bp    DNA     linear   PRI 03-FEB-2004
DEFINITION  Pongo pygmaeus UbC gene for polyubiquitin C, complete cds.
ACCESSION   AB089616
VERSION     AB089616.1  GI:28316341
KEYWORDS    .
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1
  AUTHORS   Tachikui,H., Saitou,N., Nakajima,T., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Lineage-specific homogenization of the polyubiquitin gene among
            human and great apes
  JOURNAL   J. Mol. Evol. 57 (6), 737-744 (2003)
   PUBMED   14745543
REFERENCE   2  (bases 1 to 2286)
  AUTHORS   Tachikui,H., Nakajima,T., Saito,N., Hayasaka,I., Ishida,T. and
            Inoue,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2002) Hiroshi Tachikui, Institute of Medical
            Science, University of tokyo, Division of Genetic Dignosis;
            Shirokanedai 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:tachikui@ims.u-tokyo.ac.jp, Tel:81-5449-5750,
            Fax:81-5449-5764)
FEATURES             Location/Qualifiers
     source          1. .2286
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9600"
                     /note="Orangutan with 10 ub units"
     gene            1. .2286
                     /gene="UbC"
     CDS             1. .2286
                     /gene="UbC"
                     /codon_start=1
                     /product="polyubiquitin C"
                     /protein_id="BAC56954.1"
                     /db_xref="GI:28316342"
                     /translation="MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
                     FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENV
                     KAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
                     TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKEST
                     LHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGK
                     QLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKI
                     QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
                     ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLV
                     LRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
                     GRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKE
                     GIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLE
                     VEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLR
                     GGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL
                     SDYNIQKESTLHLVLRLRGGV"
ORIGIN
```

```
  Query Match             6.2%;  Score 137;  DB 8;  Length 2286;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-65;
  Matches  137;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        456 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 1587

Qy        516 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 1647

Qy        576 GAGCCCAGTGACACCAT 592
              |||||||||||||||||
Db       1648 GAGCCCAGTGACACCAT 1664
```

```
RESULT 190
CQ700105
LOCUS       CQ700105               208 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 45031 from Patent WO02070737.
ACCESSION   CQ700105
VERSION     CQ700105.1  GI:42260069
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 45031 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .208
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             6.2%;  Score 135;  DB 6;  Length 208;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-63;
  Matches  135;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1491 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 60

Qy       1551 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 120

Qy       1611 TACAACATCCAGAAA 1625
              |||||||||||||||
Db        121 TACAACATCCAGAAA 135
```

```
RESULT 191
CS025483
LOCUS       CS025483               133 bp    DNA     linear   PAT 03-MAR-2005
DEFINITION  Sequence 67 from Patent WO2005014817.
ACCESSION   CS025483
VERSION     CS025483.1  GI:60496257
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Rotman,G., Toporoik,A., Biton,S., Savitzky,K. and Bernstein,J.
  TITLE     Chordin-like homologs
  JOURNAL   Patent: WO 2005014817-A 67 17-FEB-2005;
            Compugen Ltd. (IL); Oren, Anat (IL)
```

```
FEATURES             Location/Qualifiers
     source          1. .133
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Ubiquitin Amplicon"
ORIGIN

  Query Match             6.1%;  Score 133;  DB 6;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-62;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCA 152
              |||||||||||||
Db        121 GAATGTCAAGGCA 133


RESULT 192
CQ689709
LOCUS       CQ689709                374 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34635 from Patent WO02070737.
ACCESSION   CQ689709
VERSION     CQ689709.1  GI:42224612
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 34635 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             6.0%;  Score 132;  DB 6;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-62;
  Matches  132;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        178 CTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGT 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 CTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGT 93

Qy        238 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 153

Qy        298 TGCAAATCTTCG 309
              ||||||||||||
Db        154 TGCAAATCTTCG 165


RESULT 193
BV168033
LOCUS       BV168033                251 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm7091 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV168033
VERSION     BV168033.1  GI:48001726
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 251)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 251.
FEATURES             Location/Qualifiers
     source          1. .251
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>251
ORIGIN

  Query Match             5.7%;  Score 126;  DB 10;  Length 251;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 185

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 245

Qy        186 CAGCAG 191
              ||||||
Db        246 CAGCAG 251


RESULT 194
BD225145
LOCUS       BD225145                369 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein.
ACCESSION   BD225145
VERSION     BD225145.1  GI:33034915
KEYWORDS    JP 2002528519-A/6.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 369)
  AUTHORS   Pallente,D., Massimo,A.M.D. and Desantis,R.
  TITLE     Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein
  JOURNAL   Patent: JP 2002528519-A 6 03-SEP-2002;
            MENARINI RICERCHE SPA
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002528519-A/6
            PD   03-SEP-2002
            PF   18-OCT-1999 JP 2000579265
            PR   30-OCT-1998 IT  MI98A002330
```

```
            PI   DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
            A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
            A61K37/02,
            PC   C12N15/00
            CC   Medicinal composition having antitumor effect and containing
            CC           DNA encoding
            CC   antigenic protein
            FH   Key             Location/Qualifiers
            FT   source          1. .369
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .369
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 126;  DB 6;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 195
BD225146
LOCUS       BD225146                579 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein.
ACCESSION   BD225146
VERSION     BD225146.1  GI:33034916
KEYWORDS    JP 2002528519-A/7.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Pallente,D., Massimo,A.M.D. and Desantis,R.
  TITLE     Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein
  JOURNAL   Patent: JP 2002528519-A 7 03-SEP-2002;
            MENARINI RICERCHE SPA
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002528519-A/7
            PD   03-SEP-2002
            PF   18-OCT-1999 JP 2000579265
            PR   30-OCT-1998 IT  MI98A002330
            PI   DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
            A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
            A61K37/02,
            PC   C12N15/00
            CC   Medicinal composition having antitumor effect and containing
            CC           DNA encoding
            CC   antigenic protein
            FH   Key             Location/Qualifiers
            FT   source          1. .579
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 126;  DB 6;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 196
BD225149
LOCUS       BD225149                651 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein.
ACCESSION   BD225149
VERSION     BD225149.1  GI:33034919
KEYWORDS    JP 2002528519-A/10.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   Pallente,D., Massimo,A.M.D. and Desantis,R.
  TITLE     Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein
  JOURNAL   Patent: JP 2002528519-A 10 03-SEP-2002;
            MENARINI RICERCHE SPA
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002528519-A/10
            PD   03-SEP-2002
            PF   18-OCT-1999 JP 2000579265
            PR   30-OCT-1998 IT  MI98A002330
            PI   DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
            A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
            A61K37/02,
            PC   C12N15/00
            CC   Medicinal composition having antitumor effect and containing
            CC           DNA encoding
            CC   antigenic protein
            FH   Key             Location/Qualifiers
            FT   source          1. .651
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 126;  DB 6;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
```

```
                ||||||
Db        121 CAGAGG 126


RESULT 197
BD225147
LOCUS       BD225147                 891 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein.
ACCESSION   BD225147
VERSION     BD225147.1  GI:33034917
KEYWORDS    JP 2002528519-A/8.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 891)
  AUTHORS   Pallente,D., Massimo,A.M.D. and Desantis,R.
  TITLE     Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein
  JOURNAL   Patent: JP 2002528519-A 8 03-SEP-2002;
            MENARINI RICERCHE SPA
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002528519-A/8
            PD   03-SEP-2002
            PF   18-OCT-1999 JP 2000579265
            PR   30-OCT-1998 IT  MI98A002330
            PI   DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
            A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
            A61K37/02,
            PC   C12N15/00
            CC   Medicinal composition having antitumor effect and containing
            CC            DNA encoding
            CC   antigenic protein
            FH   Key             Location/Qualifiers
            FT   source          1. .891
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .891
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 126;  DB 6;  Length 891;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
                ||||||
Db        121 CAGAGG 126


RESULT 198
BD225148
LOCUS       BD225148                1020 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein.
ACCESSION   BD225148
VERSION     BD225148.1  GI:33034918
KEYWORDS    JP 2002528519-A/9.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1020)
  AUTHORS   Pallente,D., Massimo,A.M.D. and Desantis,R.
  TITLE     Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein
  JOURNAL   Patent: JP 2002528519-A 9 03-SEP-2002;
            MENARINI RICERCHE SPA
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002528519-A/9
            PD   03-SEP-2002
            PF   18-OCT-1999 JP 2000579265
            PR   30-OCT-1998 IT  MI98A002330
            PI   DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
            A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
            A61K37/02,
            PC   C12N15/00
            CC   Medicinal composition having antitumor effect and containing
            CC            DNA encoding
            CC   antigenic protein
            FH   Key             Location/Qualifiers
            FT   source          1. .1020
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1020
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 126;  DB 6;  Length 1020;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
                ||||||
Db        121 CAGAGG 126


RESULT 199
BD225150
LOCUS       BD225150                1737 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein.
ACCESSION   BD225150
VERSION     BD225150.1  GI:33034920
KEYWORDS    JP 2002528519-A/11.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1737)
  AUTHORS   Pallente,D., Massimo,A.M.D. and Desantis,R.
  TITLE     Medicinal composition having antitumor effect and containing DNA
            encoding antigenic protein
  JOURNAL   Patent: JP 2002528519-A 11 03-SEP-2002;
            MENARINI RICERCHE SPA
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002528519-A/11
            PD   03-SEP-2002
            PF   18-OCT-1999 JP 2000579265
            PR   30-OCT-1998 IT  MI98A002330
            PI   DINO PALLENTE,ANNA MARIA D MASSIMO,RITA DESANTIS PC
            A61K38/00,A61K35/76,A61K39/00,A61K48/00,A61P35/00,C12N15/09, PC
```

```
            A61K37/02,
            PC   C12N15/00
            CC   Medicinal composition having antitumor effect and containing
            CC           DNA encoding
            CC   antigenic protein
            FH   Key             Location/Qualifiers
            FT   source          1. .1737
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1737
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 126;  DB 6;  Length 1737;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-58;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 200
CQ659465
LOCUS       CQ659465                 168 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 4391 from Patent WO02070737.
ACCESSION   CQ659465
VERSION     CQ659465.1  GI:42127880
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 4391 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .168
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 125;  DB 6;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-58;
  Matches  125;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1423 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 1482
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 60

Qy       1483 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 1542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 120

Qy       1543 TCCCT 1547
             |||||
Db        121 TCCCT 125


RESULT 201
AX184690
LOCUS       AX184690                 255 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 385 from Patent WO0142467.
ACCESSION   AX184690
VERSION     AX184690.1  GI:15136052
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 385 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .255
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.7%;  Score 125;  DB 6;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-58;
  Matches  125;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 79

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 139

Qy        870 CAGCA 874
              |||||
Db        140 CAGCA 144


RESULT 202
BV207847
LOCUS       BV207847                 201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm224696 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV207847
VERSION     BV207847.1  GI:48178453
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
```

```
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             5.7%;  Score 124;  DB 10;  Length 201;
  Best Local Similarity   99.4%;  Pred. No. 1.8e-57;
  Matches  174;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1759 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 1818
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 60

Qy       1819 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 1878
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         61 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCYACCCTGCACCTGGTGCTCC 120

Qy       1879 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1933
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 175


RESULT 203
BD264714/c
LOCUS       BD264714                636 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compositions and methods for the therapy and diagnosis of ovarian
            cancer.
ACCESSION   BD264714
VERSION     BD264714.1  GI:33074482
KEYWORDS    JP 2002532093-A/99.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A. and Frudakis,T.N.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
  JOURNAL   Patent: JP 2002532093-A 99 02-OCT-2002;
            CORIXA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002532093-A/99
            PD   02-OCT-2002
            PF   17-DEC-1999 JP 2000588356
            PR   17-DEC-1998 US    09/215681,17-DEC-1998 US    09/216003 PR
            23-JUN-1999 US    09/338933,24-SEP-1999 US    09/404879 PI
            JENNIFER L MITCHAM,GORDON E KING,PAUL A ALGATE,TONY N FRUDAKIS PC
            C12N15/09,A61K31/7115,A61K35/14,A61K35/76,A61K39/00,A61K39/395, PC
             A61K39/395,
            PC   A61K48/00,A61P35/00,A61P37/04,C07K14/82,C07K19/00,C12N1/15, PC
              C12N1/19,
            PC
            C12N1/21,C12N5/06,C12N5/10,C12Q1/68,G01N33/53,G01N33/53,G01N33/ PC
             566,
            PC   G01N33/574,G01N33/577,C12N15/00,C12N5/00,C12N5/00 CC
            Compositions and methods for the therapy and diagnosis of  CC
                 ovarian cancer
            FH   Key             Location/Qualifiers
            FT   source          1. .636
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 204
AR238117/c
LOCUS       AR238117                636 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 99 from patent US 6468546.
ACCESSION   AR238117
VERSION     AR238117.1  GI:27283081
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Mitcham,J.L., King,G.E. and Algate,P.A.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6468546-A 99 22-OCT-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 205
AR257658/c
LOCUS       AR257658                636 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 99 from patent US 6488931.
ACCESSION   AR257658
VERSION     AR257658.1  GI:27307733
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6488931-A 99 03-DEC-2002;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="unknown"
                     /mol_type="genomic DNA"
```

```
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 206
AR283704/c
LOCUS       AR283704                636 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 99 from patent US 6528253.
ACCESSION   AR283704
VERSION     AR283704.1  GI:29720601
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for diagnosis of ovarian cancer
  JOURNAL   Patent: US 6528253-A 99 04-MAR-2003;
            Corixa Corporation; Seattle, WA
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 207
AR443378/c
LOCUS       AR443378                636 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 99 from patent US 6670463.
ACCESSION   AR443378
VERSION     AR443378.1  GI:42671157
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Mitcham,J.L., Frudakis,T.N. and King,G.E.
  TITLE     Compositions and methods for therapy of ovarian cancer
  JOURNAL   Patent: US 6670463-A 99 30-DEC-2003;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 208
AR478456/c
LOCUS       AR478456                636 bp    DNA     linear   PAT 14-MAY-2004
DEFINITION  Sequence 99 from patent US 6699664.
ACCESSION   AR478456
VERSION     AR478456.1  GI:47237108
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S. and Carter,D.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6699664-A 99 02-MAR-2004;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 209
AR640704/c
LOCUS       AR640704                636 bp    DNA     linear   PAT 20-APR-2005
DEFINITION  Sequence 99 from patent US 6858710.
ACCESSION   AR640704
VERSION     AR640704.1  GI:62775522
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
```

```
            Unclassified.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Bangur,C.S., Retter,M.W., Fanger,G.R. and Hill,P.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: US 6858710-A 99 22-FEB-2005;
            Corixa Corporation; Seattle, WA;
            WOX;
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 210
AX366332/c
LOCUS       AX366332                636 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 99 from Patent WO0206317.
ACCESSION   AX366332
VERSION     AX366332.1  GI:18697758
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
            Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
            Albone,E.
  TITLE     Compositions and methods for the therapy and diagnosis of ovarian
            cancer
  JOURNAL   Patent: WO 0206317-A 99 24-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-57;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 211
AX322311/c
LOCUS       AX322311                184 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 184 from Patent EP1162276.
ACCESSION   AX322311
VERSION     AX322311.1  GI:18093378
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kramer,M.D., Winter,H. and Reinartz,J.
  TITLE     Mrna molecules to be used as indicators of the functional and
            activational state of t-lymphocytes
  JOURNAL   Patent: EP 1162276-A 184 12-DEC-2001;
            Lynx Therapeutics GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .184
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.5%;  Score 121;  DB 6;  Length 184;
  Best Local Similarity   99.4%;  Pred. No. 8.6e-56;
  Matches  171;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2021 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 2080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 125

Qy       2081 AGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAA 2140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTTTAAGTTTCCCCTTTTAA 65

Qy       2141 GGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          64 GGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 13


RESULT 212
CQ689661
LOCUS       CQ689661                341 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 34587 from Patent WO02070737.
ACCESSION   CQ689661
VERSION     CQ689661.1  GI:42224472
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 34587 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .341
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.5%;  Score 120;  DB 6;  Length 341;
  Best Local Similarity   99.4%;  Pred. No. 3.1e-55;
  Matches  170;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        218 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 277
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 60

Qy        278 GCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 337
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCGCTGGCAAGACCATCAC 120

Qy        338 CCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 171


RESULT 213
CQ683199
LOCUS       CQ683199                 112 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 28125 from Patent WO02070737.
ACCESSION   CQ683199
VERSION     CQ683199.1  GI:42205209
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 28125 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .112
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.1%;  Score 112;  DB 6;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-51;
  Matches  112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 60

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 112


RESULT 214
CQ660012
LOCUS       CQ660012                 167 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 4938 from Patent WO02070737.
ACCESSION   CQ660012
VERSION     CQ660012.1  GI:42129945
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 4938 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .167
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.0%;  Score 109;  DB 6;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-49;
  Matches  109;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2084 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 2143
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 60

Qy       2144 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 109


RESULT 215
A87049/c
LOCUS       A87049                   795 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 140 from Patent WO9838306.
ACCESSION   A87049
VERSION     A87049.1  GI:6735834
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified sequences.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   Dolganov,G.
  TITLE     TRANSCRIPTS ENCODING IMMUNOMODULATORY POLYPEPTIDES
  JOURNAL   Patent: WO 9838306-A 140 03-SEP-1998;
            GENELABS TECH INC (US)
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /isolate="E67.SEQ"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             4.7%;  Score 104;  DB 6;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-46;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
                 ||||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6


RESULT 216
AR048461/c
LOCUS       AR048461                 795 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 140 from patent US 5821091.
ACCESSION   AR048461
VERSION     AR048461.1  GI:5970804
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   Dolganov,G.
  TITLE     Method of identifying activated T-cells
  JOURNAL   Patent: US 5821091-A 140 13-OCT-1998;
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             4.7%;  Score 104;  DB 6;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-46;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             |||||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6


RESULT 217
AR309783/c
LOCUS       AR309783                 795 bp    mRNA    linear   PAT 12-JUN-2003
DEFINITION  Sequence 140 from patent US 6555666.
ACCESSION   AR309783
VERSION     AR309783.1  GI:31701860
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   Dolganov,G.
  TITLE     Transcripts encoding immunomodulatory polypeptides
  JOURNAL   Patent: US 6555666-A 140 29-APR-2003;
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="unknown"
                     /mol_type="mRNA"
ORIGIN

  Query Match             4.7%;  Score 104;  DB 6;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-46;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             |||||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6


RESULT 218
CQ668311
LOCUS       CQ668311                 107 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 13237 from Patent WO02070737.
ACCESSION   CQ668311
VERSION     CQ668311.1  GI:42155916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 13237 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             4.6%;  Score 100;  DB 6;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 2149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 67

Qy       2150 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATT 2189
             |||||||||||||||||||||||||||||||||||||||||
Db         68 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATT 107


RESULT 219
BV207050
LOCUS       BV207050                 201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm222603 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV207050
VERSION     BV207050.1  GI:48177413
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             4.6%;  Score 100;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        531 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 161

Qy        591 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||
Db        162 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 201


RESULT 220
BV207845
LOCUS       BV207845                 201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm224693 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV207845
VERSION     BV207845.1  GI:48178450
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
```

```
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             4.6%;  Score 100;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 161

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 1617
             ||||||||||||||||||||||||||||||||||||||||
Db        162 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 201


RESULT 221
BV207848
LOCUS       BV207848                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm224704 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV207848
VERSION     BV207848.1  GI:48178454
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             4.6%;  Score 100;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1535 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 60

Qy       1595 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1634
             ||||||||||||||||||||||||||||||||||||||||
Db         61 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 100


RESULT 222
BV208263
LOCUS       BV208263                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm225974 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV208263
VERSION     BV208263.1  GI:48179002
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             4.6%;  Score 100;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1804 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 161

Qy       1864 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1903
             |||||||||||||||||||||||||||||||||||||||
Db        162 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 201


RESULT 223
BV208264
LOCUS       BV208264                201 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm225975 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
```

```
ACCESSION   BV208264
VERSION     BV208264.1  GI:48179003
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 201)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 201.
FEATURES             Location/Qualifiers
     source          1. .201
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>201
ORIGIN

  Query Match             4.6%;  Score 100;  DB 10;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1824 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 161

Qy       1884 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1923
             |||||||||||||||||||||||||||||||||||||||
Db        162 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 201


Search completed: March  6, 2006, 23:50:30
Job time : 10759 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 20:19:34 ; Search time 1222 Seconds
(without alignments)
11954.997 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 100

Total number of hits satisfying chosen parameters: 197

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2192	100.0	2192	2	AAT62351	Aat62351 Human pol
2	2192	100.0	2192	2	AAT64489	Poly-Ubiq
3	2192	100.0	2192	6	ABK83453	Abk83453 Human cDN
4	2192	100.0	2192	6	ABV78062	Abv78062 Hypoxia-r
5	2192	100.0	2192	10	ADD14753	Add14753 Human src
6	2192	100.0	2192	12	ADP02960	Adp02960 Human hou
7	2192	100.0	2192	13	ADS88458	Ads88458 Human hou
8	2192	100.0	2192	13	ADU60040	Adu60040 Housekeep
9	2192	100.0	2192	14	ADZ67404	Adz67404 Different
10	2177	99.3	2201	14	ADW02235	Adw02235 Human ubi
11	2177	99.3	2210	14	ADW02234	Adw02234 Human ubi
12	2128	97.1	3055	13	ADQ84521	Adq84521 Human tum
13	2128	97.1	3055	13	ACN39601	Acn39601 Tumour-as
14	1284	58.6	2291	10	ABT17058	Abt17058 Human MP2
15	1283	58.5	2058	14	ADV42626	Adv42626 Human psy
16	1283	58.5	2309	6	ABL69377	Abl69377 Prostate
17	1283	58.5	2309	6	ABK83973	Abk83973 Human cDN
18	1283	58.5	2309	14	ADX05493	Adx05493 Cyclin-de
19	824	37.6	824	8	ABX10337	Abx10337 DNA encod
20	719	32.8	1300	10	ADA53652	Ada53652 Human cod
21	476	21.7	693	6	ABQ55963	Abq55963 Human ova
22	449	20.5	547	2	AAZ08313	Aaz08313 Human lun
23	449	20.5	547	3	AAC79165	Aac79165 Human lun
24	449	20.5	547	4	AAD23240	Aad23240 Human lun
25	449	20.5	547	10	ADD66479	Add66479 Human lun
26	449	20.5	547	10	ADE87733	Ade87733 Human lun
27	433	19.8	460	3	AAA77883	Aaa77883 cDNA enco
28	433	19.8	460	4	AAI28621	Aai28621 Colon tum
29	433	19.8	460	8	ABZ32807	Abz32807 Human col
30	430	19.6	494	3	AAC79233	Aac79233 Human lun
31	430	19.6	494	4	AAD23309	Aad23309 Human lun
32	430	19.6	494	10	ADD66583	Add66583 Human lun
33	430	19.6	494	10	ADE87837	Ade87837 Human lun
34	426	19.4	483	2	AAZ07244	Aaz07244 Human lun
35	426	19.4	483	3	AAC79135	Aac79135 Human lun
36	426	19.4	483	4	AAD23210	Aad23210 Human lun
37	426	19.4	483	10	ADD66449	Add66449 Human lun
38	426	19.4	483	10	ADE87703	Ade87703 Human lun
39	411	18.8	412	10	ADD67000	Add67000 Human lun
40	411	18.8	412	10	ADE88254	Ade88254 Human lun
41	410	18.7	434	3	AAC79206	Aac79206 Human lun
42	410	18.7	434	4	AAD23282	Aad23282 Human lun
43	410	18.7	434	10	ADD66556	Add66556 Human lun
44	410	18.7	434	10	ADE87810	Ade87810 Human lun
45	408	18.6	442	9	ACH36460	Ach36460 Human end
c 46	407	18.6	743	14	ACL55035	Acl55035 Human col
47	403	18.4	425	13	ACN39602	Acn39602 Tumour-as
48	398	18.2	1187	3	AAF21656	Aaf21656 Human bre
49	376	17.2	389	9	ACH19114	Ach19114 Human adu
50	376	17.2	506	9	ACH33450	Ach33450 Human end
51	370	16.9	518	3	AAA69842	Aaa69842 Human ova
52	370	16.9	518	6	ABN72736	Abn72736 Ovarian c
53	370	16.9	518	9	ADA08901	Ada08901 Human ova
54	370	16.9	518	10	ADF08648	Adf08648 cDNA enco
55	370	16.9	518	10	ADG46396	Adg46396 Human ova
56	370	16.9	518	14	ADX17531	Adx17531 Human ova
c 57	307	14.0	419	13	ACF88710	Acf88710 Human SIR
58	303	13.8	400	10	ACD92815	Acd92815 Human col
59	303	13.8	412	4	AAI29110	Aai29110 Colon tum
60	303	13.8	412	8	ABZ33296	Abz33296 Human col
61	300	13.7	482	9	ACH22349	Ach22349 Human adu
62	294	13.4	393	9	ACH16408	Ach16408 Human adu
63	275	12.5	399	9	ACH20515	Ach20515 Human adu
64	270	12.3	284	2	AAX23500	Aax23500 Human neu
65	260	11.9	411	9	ACH20853	Ach20853 Human adu
66	255	11.6	543	13	ADU14740	Adu14740 Solid tum
67	254	11.6	565	13	ADU14804	Adu14804 Solid tum
68	248	11.3	585	3	AAA69735	Aaa69735 Human ova
69	248	11.3	585	6	ABN72629	Abn72629 Ovarian c
70	248	11.3	585	9	ADA08794	Ada08794 Human ova
71	248	11.3	585	10	ADF08541	Adf08541 cDNA enco
72	248	11.3	585	10	ADG46289	Adg46289 Human ova
73	248	11.3	585	14	ADX17424	Adx17424 Human ova
c 74	244	11.1	467	5	ABV45876	Abv45876 Human pro
75	243	11.1	558	13	ACF88709	Acf88709 Human SIR
76	242	11.0	286	4	AAH71976	Aah71976 Human cer
77	242	11.0	487	4	AAH70916	Aah70916 Human cer
78	228	10.4	228	8	ABZ20057	Abz20057 Group III
79	228	10.4	228	8	ABZ19866	Abz19866 Group III
c 80	228	10.4	228	8	ABZ19086	Abz19086 Group III
81	228	10.4	228	8	ABZ18791	Abz18791 Group III
82	228	10.4	228	8	ABZ19457	Abz19457 Group III
83	228	10.4	228	8	ABZ19715	Abz19715 Group III
84	228	10.4	228	13	ADR12560	Adr12560 Gene vacc
85	228	10.4	228	13	ADR12558	Adr12558 Gene vacc
86	228	10.4	228	13	ADR12564	Adr12564 Gene vacc
87	228	10.4	228	13	ADS87013	Ads87013 Human gen
88	228	10.4	228	13	ADS87017	Ads87017 Human gen
89	228	10.4	228	13	ADS87011	Ads87011 Human gen
90	211	9.6	440	6	ABT07115	Abt07115 Human ova
91	211	9.6	440	8	ABX72993	Abx72993 Human ova
92	209	9.5	228	13	ADR12562	Adr12562 Gene vacc

93	209	9.5	228	13	ADS87015	Ads87015 Human gen
94	209	9.5	7648	4	AAH48763	Aah48763 Vector pH
95	207	9.4	228	6	ABK69710	Abk69710 Human ubi
96	207	9.4	228	10	ADB67586	Adb67586 Human ubi
97	207	9.4	228	10	ADG75008	Adg75008 Human ubi
98	207	9.4	228	10	ADJ80819	Adj80819 Wilm's tu
99	207	9.4	228	12	ADJ83739	Adj83739 Human ubi
100	207	9.4	228	12	ADL57627	Adl57627 Human ubi
101	207	9.4	228	12	ADO09182	Ado09182 Human ubi
102	207	9.4	228	13	ADR12534	Adr12534 Gene vacc
103	207	9.4	228	13	ADS86985	Ads86985 Human gen
104	207	9.4	231	8	AAL51106	Aal51106 Human ubi
c 105	207	9.4	609	3	AAA16164	Aaa16164 Human col
106	207	9.4	1581	13	ADR23183	Adr23183 Yellow fl
107	205	9.4	332	3	AAC98259	Aac98259 Human col
108	203	9.3	480	10	ACF04785	Acf04785 Vector pG
109	203	9.3	504	10	ACF04786	Acf04786 Vector pG
110	201	9.2	471	9	ACH34906	Ach34906 Human end
111	201	9.2	549	13	ADU14717	Adu14717 Solid tum
112	195	8.9	234	13	ADR12570	Adr12570 Gene vacc
113	195	8.9	234	13	ADS87023	Ads87023 Human gen
c 114	191	8.7	605	3	AAA69785	Aaa69785 Human ova
c 115	191	8.7	605	6	ABN72679	Abn72679 Ovarian c
c 116	191	8.7	605	9	ADA08844	Ada08844 Human ova
c 117	191	8.7	605	10	ADF08591	Adf08591 cDNA enco
c 118	191	8.7	605	10	ADG46339	Adg46339 Human ova
c 119	191	8.7	605	14	ADX17474	Adx17474 Human ova
120	191	8.7	697	3	AAA69790	Aaa69790 Human ova
121	191	8.7	697	6	ABN72684	Abn72684 Ovarian c
122	191	8.7	697	9	ADA08849	Ada08849 Human ova
123	191	8.7	697	10	ADF08596	Adf08596 cDNA enco
124	191	8.7	697	10	ADG46344	Adg46344 Human ova
125	191	8.7	697	14	ADX17479	Adx17479 Human ova
126	188	8.6	6340	4	AAD13071	Aad13071 pcDNA3-1X
127	188	8.6	6411	4	AAD13063	Aad13063 pcDNA3-1X
128	188	8.6	6411	4	AAD13067	Aad13067 pcDNA3-Ub
129	188	8.6	6436	4	AAD13078	Aad13078 pcDNA3-1X
130	188	8.6	6459	4	AAD13087	Aad13087 pcDNA3-1X
131	188	8.6	6459	4	AAD13093	Aad13093 pcDNA3-1X
132	188	8.6	6607	4	AAD13072	Aad13072 pcDNA3-2X
133	188	8.6	6678	4	AAD13064	Aad13064 pcDNA3-2X
134	188	8.6	6703	4	AAD13079	Aad13079 pcDNA3-2X
135	188	8.6	6726	4	AAD13088	Aad13088 pcDNA3-2X
136	188	8.6	6727	4	AAD13094	Aad13094 pcDNA3-2X
137	188	8.6	6850	4	AAD13073	Aad13073 pcDNA3-3X
138	188	8.6	6921	4	AAD13065	Aad13065 pcDNA3-3X
139	188	8.6	6946	4	AAD13080	Aad13080 pcDNA3-3X
140	188	8.6	6967	4	AAD13095	Aad13095 pcDNA3-3X
141	188	8.6	6969	4	AAD13089	Aad13089 pcDNA3-3X
142	188	8.6	7051	4	AAD13110	Aad13110 pcDNA3-Ub
143	188	8.6	7052	4	AAD13111	Aad13111 pcDNA3-Ub
144	188	8.6	7053	4	AAD13105	Aad13105 pcDNA3-Ub
145	188	8.6	7093	4	AAD13074	Aad13074 pcDNA3-4X
146	188	8.6	7164	4	AAD13066	Aad13066 pcDNA3-4X
147	188	8.6	7189	4	AAD13081	Aad13081 pcDNA3-4X
148	188	8.6	7212	4	AAD13096	Aad13096 pcDNA3-4X
149	188	8.6	7212	4	AAD13090	Aad13090 pcDNA3-4X
150	188	8.6	7247	4	AAD13084	Aad13084 pcDNA3-Ub
151	188	8.6	7511	4	AAD13112	Aad13112 pcDNA3-Ub
152	188	8.6	7563	4	AAD13108	Aad13108 pcDNA3-3X
153	188	8.6	7563	4	AAD13104	Aad13104 pcDNA3-3X
154	188	8.6	7564	4	AAD13109	Aad13109 pcDNA3-3X
155	188	8.6	8022	4	AAD13100	Aad13100 pcDNA3-3X
156	179	8.2	471	8	ABZ19630	Abz19630 Group III
157	176	8.0	227	4	AAS24867	Aas24867 Human ova
158	176	8.0	335	5	AAH83506	Aah83506 Human ova
159	171	7.8	176	8	ABZ19396	Abz19396 Group III
c 160	171	7.8	294	5	ABV16078	Abv16078 Human pro
161	167	7.6	228	4	AAD13042	Aad13042 Human ubi
162	167	7.6	228	4	AAD13057	Aad13057 DNA encod
163	156	7.1	1581	13	ADR23185	Adr23185 Green flu
c 164	152	6.9	509	6	ABQ59438	Abq59438 Human col
c 165	150	6.8	150	8	ABZ19236	Abz19236 Group III
166	150	6.8	176	2	AAT19287	Aat19287 Human gen
167	147	6.7	228	13	ADR12556	Adr12556 Gene vacc
168	147	6.7	228	13	ADS87009	Ads87009 Human gen
169	141	6.4	794	6	ABQ55128	Abq55128 Human ova
c 170	137	6.2	240	4	AAS25026	Aas25026 Human ova
c 171	137	6.2	240	5	AAH83671	Aah83671 Human ova
172	135	6.2	177	10	ADL24638	Adl24638 Intestina
c 173	134	6.1	540	10	ACD92298	Acd92298 Human col
174	133	6.1	133	14	ADX55963	Adx55963 Human ubi
175	133	6.1	133	14	AEB53761	Aeb53761 Human ubi
176	133	6.1	133	14	AEC04911	Aec04911 Human UBC
177	133	6.1	133	14	AEB63482	Aeb63482 Human ubi
178	133	6.1	133	14	AEB70674	Aeb70674 Endometri
179	126	5.7	228	13	ADR12568	Adr12568 Gene vacc
180	126	5.7	228	13	ADS87021	Ads87021 Human gen
181	126	5.7	369	3	AAD00389	Aad00389 Ubiquitin
182	126	5.7	579	3	AAD00390	Aad00390 Ubiquitin
183	126	5.7	651	3	AAD00393	Aad00393 Ubiquitin
184	126	5.7	891	3	AAD00391	Aad00391 Ubiquitin
185	126	5.7	1020	3	AAD00392	Aad00392 Ubiquitin
186	126	5.7	1737	3	AAD00394	Aad00394 Ubiquitin
187	125	5.7	255	4	AAH69111	Aah69111 Human cer
c 188	123	5.6	636	3	AAA69789	Aaa69789 Human ova
c 189	123	5.6	636	6	ABN72683	Abn72683 Ovarian c
c 190	123	5.6	636	9	ADA08848	Ada08848 Human ova
c 191	123	5.6	636	10	ADF08595	Adf08595 cDNA enco
c 192	123	5.6	636	10	ADG46343	Adg46343 Human ova
c 193	123	5.6	636	14	ADX17478	Adx17478 Human ova
c 194	121	5.5	184	6	AAI69048	Aai69048 Activated
195	118	5.4	469	9	ACH41117	Ach41117 Human foe
c 196	104	4.7	795	2	AAV60065	Aav60065 Nucleic a
197	100	4.6	520	6	ABQ59388	Abq59388 Human col

ALIGNMENTS

```
RESULT 1
AAT62351
ID   AAT62351 standard; cDNA to mRNA; 2192 BP.
XX
AC   AAT62351;
XX
DT   19-JUN-1997  (first entry)
XX
DE   Human poly-ubiquitin cDNA sequence.
XX
KW   Human; ubiquitin; lymphoma; in vitro translation; diagnosis; disease;
KW   rabbit reticulocyte lysate; metabolism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             69. .2126
FT                   /*tag= a
FT                   /product= "human poly-ubiquitin protein"
FT                   /note= "CDS is claimed in Claim 1"
XX
PN   JP09000263-A.
XX
PD   07-JAN-1997.
XX
PF   21-JUN-1995;   95JP-00154616.
XX
PR   21-JUN-1995;   95JP-00154616.
XX
PA   (SAGA ) SAGAMI CHEM RES CENTRE.
XX
DR   WPI; 1997-112846/11.
DR   P-PSDB; AAW14134.
XX
PT   Human poly-ubiquitin cDNA - for prodn. of labelled poly-ubiquitin, useful
PT   for diagnosing diseases resulting from abnormalities of ubiquitin
```

PT metabolism.
XX
PS Claim 2; Page 8-10; 11pp; Japanese.
XX
CC This nucleotide sequence encodes the human poly-ubiquitin protein. The
CC gene was isolated from a cDNA library generated from polyA+ purified from
CC the human lymphoma cell line U937. Confirmation that the sequence encoded
CC the correct protein was done by in vitro translation in a rabbit
CC reticulocyte lysate. The DNA and protein can be used for diagnosis of
CC certain diseases resulting from abnormalities in metabolism of ubiquitins
CC or as a reagent for study of the ubiquitin system
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 2; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
Db    1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy   61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Db   61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy  121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Db  121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy  181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Db  181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy  241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Db  241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy  301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db  301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy  361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db  361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy  421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db  421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy  481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db  481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy  541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db  541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy  601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db  601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy  661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db  661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy  721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db  721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy  781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db  781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy  841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db  841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy  901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db  901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy  961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db  961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy 1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db 1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy 1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db 1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy 1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db 1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy 1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db 1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy 1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db 1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy 1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db 1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy 1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db 1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy 1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db 1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy 1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db 1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy 1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db 1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy 1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db 1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy 1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db 1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy 1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db 1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy 1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db 1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy 1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db 1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
```

```
Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 2
AAT64489
ID AAT64489 standard; cDNA to mRNA; 2192 BP.
XX
AC AAT64489;
XX
DT 03-JUN-1997 (first entry)
XX
DE Poly-Ubiquitin cDNA.
XX
KW Human; ubiquitin-combining enzyme E2;
KW epidermal cancer cell line KB ATCC CRL 17; ds.
XX
OS Synthetic.
XX
PN JP09037779-A.
XX
PD 10-FEB-1997.
XX
PF 31-JUL-1995; 95JP-00194376.
XX
PR 31-JUL-1995; 95JP-00194376.
XX
PA (SAGA) SAGAMI CHEM RES CENTRE.
XX
DR WPI; 1997-173039/16.
DR P-PSDB; AAW14848.
XX
PT Human ubiquitin-combining enzyme E2 - derived from human epidermal cancer
PT cell line KB ATCC CRL 17; useful as a research reagent.
XX
PS Example; Page 7-10; 10pp; Japanese.
XX
CC This sequence encodes polyubiquitin which was used to demonstrate the
CC ubiquitin binding properties of the human ubiquitin-combining enzyme E2.
CC Human ubiquitin-combining enzyme E2 and cDNA encoding it are useful as
CC tools in biological research. The human E2 protein was isolated from
CC human epidermal cancer cell line KB ATCC CRL 17
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 2; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
```

```
Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 3
ABK83453
ID ABK83453 standard; cDNA; 2192 BP.
XX
AC ABK83453;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #24.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 24; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
```

```
Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 4
ABV78062
ID   ABV78062 standard; DNA; 2192 BP.
XX
AC   ABV78062;
XX
DT   12-NOV-2002  (first entry)
XX
DE   Hypoxia-regulated protein coding sequence #82.
XX
KW   Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW   antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
KW   hypoxia; tumourigenesis; angiogenesis; apoptosis; cancer;
KW   ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
KW   preeclapmsia; atherosclerosis; inflammatory condition; wound healing;
KW   inflammation; erythropoiesis; hair loss; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200246465-A2.
XX
PD   13-JUN-2002.
XX
PF   10-DEC-2001; 2001WO-GB005458.
XX
PR   08-DEC-2000; 2000GB-00030076.
PR   08-FEB-2001; 2001GB-00003156.
PR   25-OCT-2001; 2001GB-00025666.
XX
PA   (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI   White J,  Mundy CR,  Ward NR,  Krige D,  Kingsman SM,  Harris RA;
PI   Rayner WN;
XX
DR   WPI; 2002-627238/67.
XX
PT   Identifying a gene involved in disease for treating hypoxia-regulated
PT   conditions, comprises comparing the transcriptome/proteome of two cell
PT   types under different conditions and identifying a differentially
PT   regulated gene.
XX
PS   Claim 37; Page 393; 538pp; English.
XX
CC   The present invention relates to methods for identifying genes and
CC   proteins that are implicated in a specific disease or physiological
CC   condition. The method comprises comparing the transcriptome/proteome of a
CC   specialised cell type implicated in a disease or condition with that of a
CC   second specialised cell type, under two experimental conditions, and
CC   identifying a gene that is differentially regulated in the two
CC   specialised cell types under experimental conditions. ABV77873-ABV78116
CC   and ABP65061-ABP65257 were identified using the methods of the invention.
CC   The coding sequences and proteins are useful for treating a disease in a
CC   patient, for manufacture of a medicament for treating hypoxia-regulated
CC   conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,
CC   biological response to hypoxia conditions, or hypoxic-associated
CC   pathology in a patient. The coding sequences and proteins are also useful
CC   for monitoring the therapeutic treatment of a disease or physiological
CC   condition, such as cancer, ischaemic conditions, reperfusion injury,
CC   retinopathy, neonatal stress, preeclapmsia, atherosclerosis, inflammatory
CC   conditions, wound healing, inflammation, erythropoiesis or hair loss
XX
SQ   Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2192;  DB 6;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 5
ADD14753
ID ADD14753 standard; cDNA; 2192 BP.
XX
AC ADD14753;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human src biomarker polynucleotide SEQ ID NO:147.
XX
KW predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003062395-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001981.
XX
PR 18-JAN-2002; 2002US-0350061P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX
DR WPI; 2003-636735/60.
DR P-PSDB; ADD14151.
XX
PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
PS Claim 2; SEQ ID NO 147; 139pp; English.
XX
CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 10; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
Db       1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy      61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Db      61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy     121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Db     121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy     181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Db     181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy     241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Db     241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy     301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db     301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy     361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db     361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy     421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db     421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy     481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db     481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy     541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db     541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy     601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db     601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy     661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db     661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy     721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db     721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy     781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db     781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy     841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db     841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy     901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db     901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy     961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db     961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy    1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db    1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy    1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db    1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy    1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db    1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy    1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db    1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy    1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db    1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy    1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db    1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy    1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db    1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy    1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db    1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy    1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db    1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy    1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db    1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy    1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db    1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy    1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db    1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy    1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db    1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy    1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db    1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy    1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db    1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy    1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db    1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy    1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db    1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy    2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db    2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy    2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
```

```
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 6
ADP02960
ID   ADP02960 standard; cDNA; 2192 BP.
XX
AC   ADP02960;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Human housekeeping gene cDNA #1.
XX
KW   ss; gene; human; housekeeping gene; capture probe; DNA microarray;
KW   cancer.
XX
OS   Homo sapiens.
XX
PN   JP2004135552-A.
XX
PD   13-MAY-2004.
XX
PF   16-OCT-2002; 2002JP-00302278.
XX
PR   16-OCT-2002; 2002JP-00302278.
XX
PA   (NIGA ) NGK INSULATORS LTD.
XX
DR   WPI; 2004-404613/38.
XX
PT   Novel human housekeeping gene commonly expressed in different human
PT   tissues, useful as capture probe in DNA microarray.
XX
PS   Claim 1; SEQ ID NO 1; 940pp; Japanese.
XX
CC   The invention relates to a human housekeeping gene. The housekeeping gene
CC   is useful as a capture probe in DNA microarray. The housekeeping gene is
CC   useful as a reference with respect to the normal gene expression measured
CC   on various conditions, and an expression database of normal human tissue
CC   is useful as reference information for using a disease related gene for
CC   diagnosis or treatment. The housekeeping gene when used as capture probe,
CC   enables diagnosis of cancer. The present sequence represents a human
CC   housekeeping gene cDNA.
XX
SQ   Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2192;  DB 12;  Length 2192;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
```

```
Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 7
ADS88458
ID ADS88458 standard; cDNA; 2192 BP.
XX
AC ADS88458;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human housekeeping gene cDNA sequence SEQ ID NO:1.
XX
KW housekeeping gene; human; probe; hybridisation; DNA microarray;
KW disease investigation; disease diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004035785-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2002; 2002WO-JP010753.
XX
PR 16-OCT-2002; 2002WO-JP010753.
XX
PA (NIGA) NGK INSULATORS LTD.
XX
PI Aburatani H, Yamamoto S;
XX
DR WPI; 2004-357219/33.
XX
PT Human housekeeping and tissue-specific genes, gene sets and transcription
PT products, probes and microarrays derived from them for reference use in
PT investigation and diagnosis of disease.
XX
PS Claim 5; SEQ ID NO 1; 372pp; Japanese.
XX
CC The present invention describes 1189 housekeeping genes and gene sets
CC containing them, expressed in 35 different human tissues. Also described:
CC (1) RNA and cDNA transcription products of the housekeeping genes; (2)
CC oligonucleotide probes hybridising to the housekeeping genes; (3) DNA
CC microarrays containing the oligonucleotide probes; (4) 1704 genes
CC expressed specifically in particular tissues; (5) RNA and cDNA
CC transcription products of the tissue-specific genes; (6) oligonucleotide
CC probes hybridising to the tissue-specific genes; and (7) DNA microarrays
CC containing the oligonucleotide probes. The tissue-specific genes are
CC expressed specifically in one of brain, thymus, pituitary, thyroid,
CC trachea, lung, breast, skin, skeletal muscle, heart, liver, spleen,
CC kidney, adrenal gland, pancreas, stomach, small intestine, colon,
CC bladder, prostate gland, testis, ovary, placenta, uterus, bone marrow,
CC foetal brain or foetal liver. The housekeeping genes can be used as
CC reference genes for the investigation and diagnosis of disease. The
CC present sequence represents a human housekeeping gene cDNA sequence which
CC is given in the exemplification of the present invention.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 13; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

```
RESULT 8
ADU60040
ID   ADU60040 standard; cDNA; 2192 BP.
XX
AC   ADU60040;
XX
DT   10-FEB-2005  (first entry)
XX
DE   Housekeeping gene cDNA, SEQ ID 1.
XX
KW   Housekeeping; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2004229233-A1.
XX
PD   18-NOV-2004.
XX
PF   15-OCT-2003; 2003US-00684422.
XX
PR   16-OCT-2002; 2002US-0418614P.
XX
PA   (NIGA ) NGK INSULATORS LTD.
XX
PI   Aburatani H,  Yamamoto S;
XX
DR   WPI; 2004-813237/80.
```

XX
PT New human housekeeping genes and tissue-specific genes, used as a
PT reference for a normal or disease-associated gene expression measured
PT under various conditions.
XX
PS Claim 5; SEQ ID NO 1; 93pp; English.
XX
CC The present invention relates to novel human housekeeping genes and
CC tissue-specific genes. The present sequence is a transcription product of
CC one such gene. Also claimed is a DNA microarray carrying the set of
CC transcription products of the genes of the invention. The housekeeping
CC genes and tissue-specific genes are useful as references for a normal
CC gene expression or disease-associated gene expression measured under
CC various conditions. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC from seqdata.uspto.gov/sequence.html?DocID=20040229233.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 13; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
```

```
Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 9
ADZ67404
ID ADZ67404 standard; cDNA; 2192 BP.
XX
AC ADZ67404;
XX
DT 14-JUL-2005 (first entry)
XX
DE Differentially expressed gene #5 in breast cancer relapse risk patient.
XX
KW prognosis; breast tumor; cancer; gene expression; ss.
XX
OS Homo sapiens.
XX
PN FR2861744-A1.
XX
PD 06-MAY-2005.
XX
PF 05-NOV-2003; 2003FR-00012966.
XX
PR 05-NOV-2003; 2003FR-00012966.
XX
PA (INMR) BIOMERIEUX SA.
PA (BERA-) CENT BERARD LEON.
XX
PI Lacroix B, Krause A, Puisieux A, Bachelot T;
XX
DR WPI; 2005-348176/36.
XX
PT Prognosis of breast cancer, useful for identifying patients at high risk
PT of relapse, based on analysis of differential expression of specific
PT genes.
XX
PS Claim 1; SEQ ID NO 5; 154pp; French.
XX
CC The invention relates to a method for prognosis of breast cancer
CC comprising extracting biological material (A) from a patient sample;
CC contacting (A) with a reagent (R) specific for target genes (TG) having
CC any of 68 sequences reproduced; and determining expression of TG. The
CC method is used to identify patients at high risk of relapse, particularly
CC to assist selection of primary treatment. This sequence corresponds to a
CC differentially regulated gene found in patients at risk of a relapse into
CC breast cancer.
XX
SQ Sequence 2192 BP; 598 A; 574 C; 574 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 2192; DB 14; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 10
ADW02235
ID ADW02235 standard; cDNA; 2201 BP.
XX
AC ADW02235;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human ubiquitin C cDNA sequence.
XX
KW DNA detection; breast cancer; ubiquitin C; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2005001135-A2.
XX
PD 06-JAN-2005.
XX
PF 02-JUN-2004; 2004WO-US017572.
XX
PR 03-JUN-2003; 2003US-0475812P.
PR 30-JAN-2004; 2004US-00769476.
XX
PA (ARCT-) ARCTURUS BIOSCIENCE INC.
XX
PI Elander MG, Ma X, Baer TM;
XX
DR WPI; 2005-075586/08.
DR GENBANK; BC000449.
XX
PT New microarray, useful for studying functional gene networks and gene
PT expression as the result of a disease state in a cell, tissue, organ, or
PT subject such as cells that are pre-cancerous, cancerous, or invasive such
PT as breast cancer.
XX
PS Example 2; Page 35; 45pp; English.
XX
CC The invention relates to a microarray comprising at least 5
CC oligonucleotide probes, each of 150 nucleotides or less in length, and
CC complementary to at least 10 consecutive nucleotides of an mRNA molecule,
CC where the 10 consecutive nucleotides is, in its entirety, less than 360
CC nucleotides from the site of poly(A) addition of the mRNA molecule.
CC Analyzing gene expression in a cell comprises preparing a polynucleotide
CC comprising gene sequences expressed in the cell and hybridizing the
CC polynucleotide to a microarray above. The cell is from an FFPE sample,
CC preferably a breast cancer cell. The cell has been isolated by
CC microdissection of a cell-containing sample. The method may comprise
CC quantitative PCR amplification of a sequence containing at least 10
CC consecutive nucleotides of an mRNA molecule present in the cell, where
CC the 10 consecutive nucleotides is, in its entirety, less than 360
CC nucleotides from the site of poly(A) addition of the mRNA molecule and
CC where the mRNA is optionally a reference sequence. The microarray is
CC useful for conducting quantitative and qualitative analysis of gene
CC expression and for studying gene expression as the result of a disease
CC state in a cell, tissue, organ, or subject, particularly the expression
CC of genes in cells that are not normal, pre-cancerous, cancerous, or
CC invasive such as breast cancer. The microarray is also useful for
CC studying functional gene networks. The methods are useful for detecting
CC cancer cells (e.g. breast cancer cells), facilitating diagnosis of cancer
CC and the severity of a cancer in a subject, facilitating a determination
CC of the prognosis of a subject, and assessing the responsiveness of the
CC subject to therapy. The present sequence represents a human ubiquitin C,
CC clone IMAGE:2821375, that can be detected using the method of the
CC invention.
XX
SQ Sequence 2201 BP; 620 A; 569 C; 570 G; 442 T; 0 U; 0 Other;

```
Query Match             99.3%;  Score 2177;  DB 14;  Length 2201;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
              |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177


RESULT 11
ADW02234
ID   ADW02234 standard; DNA; 2210 BP.
XX
AC   ADW02234;
XX
DT   24-MAR-2005  (first entry)
XX
DE   Human ubiquitin C related nucleotide sequence.
XX
KW   DNA detection; breast cancer; ubiquitin C; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005001135-A2.
XX
PD   06-JAN-2005.
XX
PF   02-JUN-2004; 2004WO-US017572.
XX
PR   03-JUN-2003; 2003US-0475812P.
PR   30-JAN-2004; 2004US-00769476.
XX
PA   (ARCT-) ARCTURUS BIOSCIENCE INC.
XX
PI   Elander MG,  Ma X,  Baer TM;
XX
DR   WPI; 2005-075586/08.
DR   GENBANK; BC000449.
XX
PT   New microarray, useful for studying functional gene networks and gene
PT   expression as the result of a disease state in a cell, tissue, organ, or
PT   subject such as cells that are pre-cancerous, cancerous, or invasive such
PT   as breast cancer.
XX
PS   Example 2; Page 34-35; 45pp; English.
XX
CC   The invention relates to a microarray comprising at least 5
CC   oligonucleotide probes, each of 150 nucleotides or less in length, and
CC   complementary to at least 10 consecutive nucleotides of an mRNA molecule,
CC   where the 10 consecutive nucleotides is, in its entirety, less than 360
CC   nucleotides from the site of poly(A) addition of the mRNA molecule.
CC   Analyzing gene expression in a cell comprises preparing a polynucleotide
CC   comprising gene sequences expressed in the cell and hybridizing the
CC   polynucleotide to a microarray above. The cell is from an FFPE sample,
CC   preferably a breast cancer cell. The cell has been isolated by
CC   microdissection of a cell-containing sample. The method may comprise
CC   quantitative PCR amplification of a sequence containing at least 10
CC   consecutive nucleotides of an mRNA molecule present in the cell, where
CC   the 10 consecutive nucleotides is, in its entirety, less than 360
CC   nucleotides from the site of poly(A) addition of the mRNA molecule and
CC   where the mRNA is optionally a reference sequence. The microarray is
CC   useful for conducting quantitative and qualitative analysis of gene
CC   expression and for studying gene expression as the result of a disease
CC   state in a cell, tissue, organ, or subject, particularly the expression
CC   of genes in cells that are not normal, pre-cancerous, cancerous, or
CC   invasive such as breast cancer. The microarray is also useful for
CC   studying functional gene networks. The methods are useful for detecting
CC   cancer cells (e.g. breast cancer cells), facilitating diagnosis of cancer
CC   and the severity of a cancer in a subject, facilitating a determination
CC   of the prognosis of a subject, and assessing the responsiveness of the
CC   subject to therapy. The present sequence represents a human nucleotide
CC   sequence that is similar to ubiquitin C, clone MGC:8448 IMAGE:2821375,
CC   that can be detected using the method of the invention.
XX
SQ   Sequence 2210 BP; 622 A; 571 C; 575 G; 442 T; 0 U; 0 Other;

  Query Match             99.3%;  Score 2177;  DB 14;  Length 2210;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2177;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        10 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 69

Qy        76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        70 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 129

Qy       136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       130 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 189

Qy       196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       190 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 249

Qy       256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       250 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 309

Qy       316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       310 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 369

Qy       376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       370 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 429

Qy       436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       430 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 489

Qy       496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       490 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 549

Qy       556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       550 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 609

Qy       616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       610 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 669

Qy       676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       670 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 729

Qy       736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       730 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 789

Qy       796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       790 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 849

Qy       856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       850 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 909

Qy       916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       910 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 969

Qy       976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       970 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1029
```

```
Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1089

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1149

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1209

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1269

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1329

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1389

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1449

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1509

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1569

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1629

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1689

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1749

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1809

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1810 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1869

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1870 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1929

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1930 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1989

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1990 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2049

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2109

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2110 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2169

Qy       2176 AAAGTTGTTGCATTCCC 2192
             |||||||||||||||||
Db       2170 AAAGTTGTTGCATTCCC 2186
```

RESULT 12
ADQ84521
ID ADQ84521 standard; cDNA; 3055 BP.
XX
AC ADQ84521;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1335.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 1335; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative

CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 3055 BP; 736 A; 744 C; 874 G; 701 T; 0 U; 0 Other;

Query Match 97.1%; Score 2128; DB 13; Length 3055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
Db       876 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 935

Qy       125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
Db       936 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 995

Qy       185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
Db       996 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 1055

Qy       245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
Db      1056 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 1115

Qy       305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
Db      1116 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 1175

Qy       365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
Db      1176 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 1235

Qy       425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
Db      1236 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 1295

Qy       485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
Db      1296 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1355

Qy       545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
Db      1356 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1415

Qy       605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
Db      1416 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 1475

Qy       665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
Db      1476 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 1535

Qy       725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
Db      1536 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1595

Qy       785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
Db      1596 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1655

Qy       845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
Db      1656 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1715

Qy       905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
Db      1716 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1775

Qy       965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
Db      1776 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1835

Qy      1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
Db      1836 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1895

Qy      1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
Db      1896 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1955

Qy      1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
Db      1956 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 2015

Qy      1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
Db      2016 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 2075

Qy      1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
Db      2076 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 2135

Qy      1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
Db      2136 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 2195

Qy      1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
Db      2196 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 2255

Qy      1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
Db      2256 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 2315

Qy      1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
Db      2316 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 2375

Qy      1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
Db      2376 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 2435

Qy      1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
Db      2436 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 2495

Qy      1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
Db      2496 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 2555

Qy      1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
Db      2556 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 2615

Qy      1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
Db      2616 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 2675

Qy      1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1924
Db      2676 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 2735

Qy      1925 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1984
Db      2736 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 2795

Qy      1985 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2044
```

```
Db       2796 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2855

Qy       2045 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2856 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2915

Qy       2105 GCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2916 GCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2975

Qy       2165 TTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||
Db       2976 TTCCTTTCAATAAAGTTGTTGCATTCCC 3003
```

```
RESULT 13
ACN39601
ID   ACN39601 standard; cDNA; 3055 BP.
XX
AC   ACN39601;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) cDNA DNA325803, SEQ ID NO:3853.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   P-PSDB; ABM81496.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 1; SEQ ID NO 3853; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT nucleic acid of the invention
XX
SQ   Sequence 3055 BP; 736 A; 744 C; 874 G; 701 T; 0 U; 0 Other;

  Query Match            97.1%;  Score 2128;  DB 13;  Length 3055;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        876 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 935

Qy        125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        936 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 995

Qy        185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        996 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 1055

Qy        245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1056 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 1115

Qy        305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1116 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 1175

Qy        365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1176 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 1235

Qy        425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1236 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 1295

Qy        485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1296 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1355

Qy        545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1356 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1415

Qy        605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1416 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 1475

Qy        665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1476 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 1535

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1536 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1595

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1596 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1655

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1656 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1715

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1716 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1775

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1776 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1835
Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
Db       1836 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1895
Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
Db       1896 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1955
Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
Db       1956 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 2015
Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
Db       2016 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 2075
Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
Db       2076 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 2135
Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
Db       2136 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 2195
Qy       1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
Db       2196 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 2255
Qy       1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
Db       2256 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 2315
Qy       1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
Db       2316 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 2375
Qy       1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
Db       2376 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 2435
Qy       1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
Db       2436 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 2495
Qy       1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
Db       2496 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 2555
Qy       1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
Db       2556 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 2615
Qy       1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
Db       2616 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 2675
Qy       1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1924
Db       2676 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 2735
Qy       1925 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1984
Db       2736 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 2795
Qy       1985 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2044
Db       2796 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2855
Qy       2045 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2104
Db       2856 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2915
Qy       2105 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2164
Db       2916 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2975
Qy       2165 TTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db       2976 TTCCTTTCAATAAAGTTGTTGCATTCCC 3003
```

RESULT 14
ABT17058
ID ABT17058 standard; DNA; 2291 BP.
XX
AC ABT17058;
XX
DT 03-APR-2003 (first entry)
XX
DE Human MP21 gene UbC SEQ ID No 32.
XX
KW Cytostatic; p21 pathway modulating agent; cancer; angiogenic; apoptotic;
KW cell proliferation disorder; MP21; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003006990-A1.
XX
PD 23-JAN-2003.
XX
PF 10-JUL-2002; 2002WO-US021549.
XX
PR 12-JUL-2001; 2001US-0305017P.
PR 10-OCT-2001; 2001US-0328491P.
PR 15-FEB-2002; 2002US-0357452P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Li D, Funke RP;
XX
DR WPI; 2003-221779/21.
DR P-PSDB; ABJ19784.
XX
PT Identifying candidate p21 pathway modulator, by contacting an assay
PT system having modifiers of p21 polypeptide or gene with a test agent to
PT provide a reference activity in system and detecting test agent-biased
PT activity.
XX
PS Example; Page 107-108; 199pp; English.
XX
CC The invention relates to a novel method for identifying a candidate p21
CC pathway modulating agent. The novel method comprises contacting an assay
CC system, comprising a purified MP21 polypeptide (modifier of p21) or
CC nucleic acid, with a test agent under conditions, so that but for the
CC presence of a test agent, the assay system provides a reference activity
CC and detection of test agent-biased activity of the assay system. The
CC novel method of the invention is useful for identifying a candidate p21
CC pathway modulating agent. The invention also includes a method for
CC modulating the p21 pathway of a cell, and a method for diagnosing a
CC disease e.g. cancer in a patient. The identified modulators are useful in
CC diagnosis, therapy and pharmaceutical development. The modulators are
CC useful in a variety of diagnostic and therapeutic applications including
CC angiogenic, apoptotic and cell proliferation disorders. This
CC polynucleotide sequence represents a gene encoding an MP21 protein of the
CC invention
XX
SQ Sequence 2291 BP; 591 A; 585 C; 616 G; 498 T; 0 U; 1 Other;

Query Match 58.6%; Score 1284; DB 10; Length 2291;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1584; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy         65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
```

```
Db        421 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 480
Qy        125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
Db        481 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 540
Qy        185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
Db        541 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 600
Qy        245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
Db        601 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 660
Qy        305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
Db        661 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCAT 720
Qy        365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
Db        721 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 780
Qy        425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
Db        781 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 840
Qy        485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
Db        841 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 900
Qy        545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
Db        901 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 960
Qy        605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
Db        961 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 1020
Qy        665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
Db       1021 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 1080
Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
Db       1081 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1140
Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
Db       1141 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1200
Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
Db       1201 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1260
Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
Db       1261 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1320
Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
Db       1321 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 1380
Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
Db       1381 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1440
Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
Db       1441 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1500
Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
Db       1501 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1560
```

```
Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
Db       1561 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1620
Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
Db       1621 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1680
Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
Db       1681 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1740
Qy       1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
Db       1741 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1800
Qy       1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
Db       1801 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCNCGAAGTGGAGCCGAGTGACACCAT 1860
Qy       1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
Db       1861 CGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1920
Qy       1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
Db       1921 GATCTTTGCCGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1980
Qy       1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
Db       1981 AGAGTCCACCCTGCACCTGGTGCTCCGTCT 2010
```

RESULT 15
ADV42626
ID ADV42626 standard; cDNA; 2058 BP.
XX
AC ADV42626;
XX
DT 10-MAR-2005 (first entry)
XX
DE Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 254.
XX
KW microarray; psychoneuroendocrinimmune; chronic fatigue;
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX
OS Homo sapiens.
XX
PN WO2004108899-A2.
XX
PD 16-DEC-2004.
XX
PF 04-JUN-2004; 2004WO-US017686.
XX
PR 04-JUN-2003; 2003US-0475915P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nicholson A, Vernon SD;
XX
DR WPI; 2005-031682/03.
XX
PT New microarray comprising probes for genes involved in
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT condition associated with PNI activity, e.g., inflammatory or infectious
PT diseases.
XX
PS Claim 1; SEQ ID NO 254; 254pp; English.
XX
CC The invention relates to a new microarray which comprises probes for
CC genes involved in psychoneuroendocrinimmune (PNI) activity. The

CC microarray is useful in diagnosing a condition associated with PNI
CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
CC cancer and infection. The present sequence represents a
CC psychoneuroendocrinimmune gene expressed sequence tag. Note the
CC specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
CC SEQ ID NO 1829 are provided.
XX
SQ Sequence 2058 BP; 574 A; 538 C; 547 G; 399 T; 0 U; 0 Other;

Query Match 58.5%; Score 1283; DB 14; Length 2058;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1483; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy       606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy       666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy       726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       202 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 261

Qy       786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
            |||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||
Db       262 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 321

Qy       846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       322 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 381

Qy       906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       382 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 441

Qy       966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       442 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 501

Qy      1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       502 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 561

Qy      1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       562 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 621

Qy      1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       622 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 681

Qy      1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       682 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 741

Qy      1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       742 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 801

Qy      1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       802 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 861

Qy      1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       862 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 921

Qy      1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        922 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 981

Qy      1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        982 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1041

Qy      1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1042 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1101

Qy      1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
            ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1102 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1161

Qy      1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1162 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1221

Qy      1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1281

Qy      1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1282 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1341

Qy      1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1401

Qy      1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1402 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1461

Qy      1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1462 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1508
```

RESULT 16
ABL69377
ID ABL69377 standard; DNA; 2309 BP.
XX
AC ABL69377;
XX
DT 15-MAY-2002 (first entry)
XX
DE Prostate cancer related gene sequence SEQ ID NO:7714.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.

```
PR   25-SEP-2000; 2000US-0234923P.
PR   25-SEP-2000; 2000US-0234924P.
PR   25-SEP-2000; 2000US-0235077P.
PR   25-SEP-2000; 2000US-0235082P.
PR   25-SEP-2000; 2000US-0235134P.
PR   25-SEP-2000; 2000US-0235280P.
PR   26-SEP-2000; 2000US-0235637P.
PR   26-SEP-2000; 2000US-0235638P.
PR   27-SEP-2000; 2000US-0235711P.
PR   27-SEP-2000; 2000US-0235720P.
PR   27-SEP-2000; 2000US-0235840P.
PR   27-SEP-2000; 2000US-0235863P.
PR   28-SEP-2000; 2000US-0236028P.
PR   28-SEP-2000; 2000US-0236032P.
PR   28-SEP-2000; 2000US-0236033P.
PR   28-SEP-2000; 2000US-0236034P.
PR   28-SEP-2000; 2000US-0236109P.
PR   28-SEP-2000; 2000US-0236111P.
PR   29-SEP-2000; 2000US-0236842P.
PR   29-SEP-2000; 2000US-0236891P.
PR   02-OCT-2000; 2000US-0237172P.
PR   02-OCT-2000; 2000US-0237173P.
PR   02-OCT-2000; 2000US-0237278P.
PR   02-OCT-2000; 2000US-0237294P.
PR   02-OCT-2000; 2000US-0237295P.
PR   02-OCT-2000; 2000US-0237316P.
PR   03-OCT-2000; 2000US-0237425P.
PR   03-OCT-2000; 2000US-0237598P.
PR   03-OCT-2000; 2000US-0237604P.
PR   03-OCT-2000; 2000US-0237606P.
PR   03-OCT-2000; 2000US-0237608P.
PR   01-NOV-2000; 2000US-0244867P.
PR   01-NOV-2000; 2000US-0245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a chemical
PT   agent to be tested for anti-neoplastic activity, and determining a change
PT   in expression of a gene of a signature gene set.
XX
PS   Claim 1; SEQ ID NO 7714; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an anti-
CC   neoplastic agent. The method involves exposing cells to a chemical agent
CC   to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening an
CC   anti-neoplastic agent, and can be used for producing a product which is
CC   the data collected with respect to the anti-neoplastic agent as a result
CC   of M1, and the data is sufficient to convey the chemical structure and/or
CC   properties of the agent. M1 can be used in the treatment of cancer such
CC   as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC   prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC   cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC   cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC   tumour
XX
SQ   Sequence 2309 BP; 624 A; 571 C; 592 G; 522 T; 0 U; 0 Other;

  Query Match             58.5%;  Score 1283;  DB 6;  Length 2309;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              |||||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||

Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516

Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576

Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636

Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696

Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756

Qy       1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816

Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876

Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936

Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996

Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056

Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116

Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176

Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236

Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||

Db       1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296
```

```
Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356

Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416

Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643
```

RESULT 17
ABK83973
ID ABK83973 standard; cDNA; 2309 BP.
XX
AC ABK83973;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #544.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 544; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2309 BP; 624 A; 571 C; 592 G; 522 T; 0 U; 0 Other;

Query Match 58.5%; Score 1283; DB 6; Length 2309;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1483; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516

Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576

Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636

Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696

Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
```

```
Db        697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756

Qy       1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816

Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876

Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936

Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996

Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056

Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116

Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176

Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236

Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296

Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356

Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416

Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643


RESULT 18
ADX05493
ID   ADX05493 standard; DNA; 2309 BP.
XX
AC   ADX05493;
XX
DT   21-APR-2005  (first entry)
XX
DE   Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 58.
XX
KW   cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005012875-A2.
XX
PD   10-FEB-2005.
XX
PF   29-JUL-2004; 2004WO-US024424.
XX
PR   29-JUL-2003; 2003US-0490890P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Li M,  Rupnow BA,  Webster KR,  Jackson DG,  Wong TW;
XX
DR   WPI; 2005-163068/17.
DR   P-PSDB; ADX05494.
XX
PT   Biomarkers useful for predicting or determining the response of a mammal
PT   to a cancer treatment comprising administration of a modulator of cyclin-
PT   dependent kinase activity.
XX
PS   Claim 5; SEQ ID NO 58; 141pp; English.
XX
CC   This invention describes a novel method of predicting or determining
CC   whether a mammal will respond or is responding to an anti-cancer agent
CC   that modulates cyclin-dependent kinase (cdk) activity. The method
CC   comprises measuring the level of one or more biomarkers selected from
CC   2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC   NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC   invention is utilized in a kit for determining or predicting whether
CC   patient would be susceptible or resistant to treatment by an agent
CC   modulating cdk activity. The invention also describes a method for
CC   utilizing individualized genetic profiles for treating diseases and
CC   disorders based on patient's response and molecular level, specialized
CC   microarrays comprising the biomarkers described, antibodies directed
CC   against the biomarkers and a cell culture model to identify biomarkers.
CC   The cdk modulator is preferably N-5-[[5-(1,1-Dimethylethyl)-2-
CC   oxazolyl]methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC   tartaric acid salt. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
CC   sequence encodes a biomarker used in the method of the invention.
XX
SQ   Sequence 2309 BP; 624 A; 571 C; 592 G; 522 T; 0 U; 0 Other;

  Query Match             58.5%;  Score 1283;  DB 14;  Length 2309;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              |||||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||
Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
```

```
Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516
Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
Db        517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576
Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
Db        577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636
Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
Db        637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696
Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
Db        697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756
Qy       1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
Db        757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816
Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
Db        817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876
Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
Db        877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936
Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
Db        937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996
Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
Db        997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056
Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
Db       1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116
Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
Db       1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176
Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
Db       1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236
Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
Db       1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296
Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
Db       1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356
Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
Db       1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416
Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
Db       1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476
Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
Db       1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536
Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
Db       1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596
Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
Db       1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643
```

RESULT 19
ABX10337
ID ABX10337 standard; DNA; 824 BP.
XX
AC ABX10337;
XX
DT 28-JAN-2003 (first entry)
XX
DE DNA encoding protein differentially regulated in prostate cancer #6.
XX
KW Prostate cancer; gene expression; differential regulation;
KW molecular marker; drug target; cancer detection; cancer diagnosis;
KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200281638-A2.
XX
PD 17-OCT-2002.
XX
PF 08-APR-2002; 2002WO-US010824.
XX
PR 06-APR-2001; 2001US-0281731P.
PR 06-APR-2001; 2001US-0281732P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Jay G;
XX
DR WPI; 2003-058520/05.
DR P-PSDB; ABU07435.
XX
PT Novel genes which are differentially regulated in prostate cancer, useful
PT for diagnosing prostate cancer in prostate tissue sample and assessing
PT therapeutic or preventive intervention in prostate cancer patients.
XX
PS Claim 1; Page 72-73; 416pp; English.
XX
CC The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity. (I)
CC is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.
CC The polypeptide encoded by (I) can be used as target for therapy or drug
CC discovery. (I) can also be used for expressing the polypeptide and thus
CC for searching specific binding partners of the polypeptide. (I) is useful

```
CC   in therapeutic applications to treat prostate cancer. The identification
CC   of specific genes, and groups of genes, expressed in pathways
CC   physiologically relevant to prostate cancer permits the definition of
CC   functional and disease pathways and the delineation of targets in these
CC   pathways which are useful in diagnostic, therapeutic, and clinical
CC   applications. This sequence encodes a protein diferentially regulated in
CC   prostate cancer
XX
SQ   Sequence 824 BP; 222 A; 215 C; 220 G; 167 T; 0 U; 0 Other;

  Query Match            37.6%;  Score 824;  DB 8;  Length 824;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  824;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1317 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 60

Qy       1377 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 120

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 180

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 240

Qy       1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 300

Qy       1617 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 360

Qy       1677 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 420

Qy       1737 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 480

Qy       1797 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 540

Qy       1857 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 600

Qy       1917 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 1976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 660

Qy       1977 AAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAA 2036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAA 720

Qy       2037 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCAC 2096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCAC 780

Qy       2097 TTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAA 2140
             ||||||||||||||||||||||||||||||||||||||||||||
Db         781 TTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAA 824

RESULT 20
ADA53652
ID   ADA53652 standard; cDNA; 1300 BP.
XX
AC   ADA53652;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Human coding sequence, SEQ ID 1220.
XX
KW   Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW   Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW   inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX
OS   Homo sapiens.
XX
PN   EP1293569-A2.
XX
PD   19-MAR-2003.
XX
PF   21-MAR-2002; 2002EP-00006586.
XX
PR   14-SEP-2001; 2001JP-00328381.
PR   24-JAN-2002; 2002US-0350435P.
XX
PA   (HELI-) HELIX RES INST.
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-395539/38.
DR   P-PSDB; ADA55291.
XX
PT   New polynucleotides encoding full-length polypeptides, e.g. secretory
PT   and/or membrane proteins, useful for developing medicines for diseases in
PT   which the gene is involved, or as target molecules for gene therapy.
XX
PS   Claim 1; SEQ ID NO 1220; 205pp; English.
XX
CC   The present invention relates to novel human secretory or membrane
CC   proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC   ADA54071). The coding sequences are useful in the gene therapy of
CC   diseases caused by abnormalities of the proteins, e.g. cancer,
CC   inflammatory diseases, osteoporosis or neurological disease.
XX
SQ   Sequence 1300 BP; 358 A; 335 C; 336 G; 271 T; 0 U; 0 Other;

  Query Match            32.8%;  Score 719;  DB 10;  Length 1300;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  769;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 316

Qy        408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 376

Qy        468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 436

Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 496

Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 556

Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        557 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 616

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 676

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 736

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 796

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 856

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 916

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        917 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 976

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1117
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1026
```

RESULT 21
ABQ55963
ID ABQ55963 standard; cDNA; 693 BP.
XX
AC ABQ55963;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HPDQC34 cDNA, SEQ ID NO:1843.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
PD 03-JAN-2002.
XX
PF 07-JUN-2001; 2001WO-US018569.
XX
PR 07-JUN-2000; 2000US-0209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-147878/19.
DR P-PSDB; ABP42886.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
PS Claim 1; SEQ ID NO 1843; 2922pp; English.
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 693 BP; 197 A; 175 C; 182 G; 128 T; 0 U; 11 Other;

Query Match 21.7%; Score 476; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 7.3e-226;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1257 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 105

Qy       1317 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 165

Qy       1377 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 225

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 285

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 345

Qy       1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 405

Qy       1617 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 465

Qy       1677 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1732
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 521
```

RESULT 22

```
AAZ08313
ID   AAZ08313 standard; cDNA; 547 BP.
XX
AC   AAZ08313;
XX
DT   13-OCT-1999  (first entry)
XX
DE   Human lung tumour protein SAL-43 5' cDNA sequence.
XX
KW   Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW   immunotherapy; detection; inhibition; ss.
XX
OS   Homo sapiens.
XX
PN   WO9938973-A2.
XX
PD   05-AUG-1999.
XX
PF   26-JAN-1999;   99WO-US001642.
XX
PR   28-JAN-1998;   98US-00015022.
PR   28-JAN-1998;   98US-00015029.
PR   18-MAR-1998;   98US-00040828.
PR   18-MAR-1998;   98US-00040831.
PR   23-JUL-1998;   98US-00122191.
PR   23-JUL-1998;   98US-00122192.
PR   22-DEC-1998;   98US-00219245.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Lodes MJ,  Frudakis TN,  Mohamath R;
XX
DR   WPI; 1999-479187/40.
DR   P-PSDB; AAY20558.
XX
PT   Lung tumor specific polynucleotides for inhibiting the development of
PT   lung cancer.
XX
PS   Claim 13; Page 143; 171pp; English.
XX
CC   The present invention describes lung tumour specific polynucleotides and
CC   tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325 represent
CC   specifically claimed polynucleotides, and AAY29486 to AAY29571 represent
CC   amino acid sequences from the present invention. The lung tumour specific
CC   polynucleotides and polypeptides can be used in pharmaceutical
CC   compositions and vaccines to inhibit the development of lung cancer. They
CC   can also be used to detect lung cancer in a patient. Probes and
CC   antibodies derived from the lung tumour sequences are useful in detection
CC   of lung cancer
XX
SQ   Sequence 547 BP; 144 A; 145 C; 146 G; 112 T; 0 U; 0 Other;

  Query Match            20.5%;  Score 449;  DB 2;  Length 547;
  Best Local Similarity  99.8%;  Pred. No. 1.9e-212;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548
              ||||||||||||||||||||
Db        528 AGATCTTCGTGAAGACCCTG 547


RESULT 23
AAC79165
ID   AAC79165 standard; cDNA; 547 BP.
XX
AC   AAC79165;
XX
DT   05-FEB-2001  (first entry)
XX
DE   Human lung tumour-specific cDNA #118.
XX
KW   Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200060077-A2.
XX
PD   12-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US008560.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Lodes MJ,  Mohamath R,  Secrist H;
XX
DR   WPI; 2000-638466/61.
DR   P-PSDB; AAB44482.
XX
PT   Novel lung tumor polypeptides and polynucleotides, useful for detecting,
PT   monitoring or treating cancer, especially lung cancer.
XX
PS   Claim 1; Page 158; 243pp; English.
XX
CC   The present sequence is given in a specification relating to compounds
CC   for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC   least an immunogenic part of a lung tumour protein are disclosed. The
CC   polypeptides are useful for inhibiting the development of cancer,
CC   especially lung cancer. Samples of T cells expressing the polypeptides
CC   may be used to inhibit the development of cancer. The polypeptides are
CC   also useful for detecting and monitoring the progression of cancer,
CC   especially lung cancer
XX
SQ   Sequence 547 BP; 144 A; 145 C; 146 G; 112 T; 0 U; 0 Other;

  Query Match            20.5%;  Score 449;  DB 3;  Length 547;
  Best Local Similarity  99.8%;  Pred. No. 1.9e-212;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
```

```
Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168

Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228

Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288

Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348

Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408

Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468

Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528

Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548

Db        528 AGATCTTCGTGAAGACCCTG 547


RESULT 24
AAD23240
ID   AAD23240 standard; cDNA; 547 BP.
XX
AC   AAD23240;
XX
DT   26-FEB-2002  (first entry)
XX
DE   Human lung tumour-specific protein SAL-43 cDNA.
XX
KW   Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW   antisense-therapy; vaccine; immune response; lung cancer; SAL-43; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             68. .547
FT                   /*tag= a
FT                   /product= "Human SAL-43 protein"
FT                   /note= "CDS does not include stop codon"
FT                   /partial
XX
PN   WO200172295-A2.
XX
PD   04-OCT-2001.
XX
PF   28-MAR-2001; 2001WO-US009991.
XX
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   22-SEP-2000; 2000US-0234517P.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI   Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
DR   WPI; 2001-639201/73.
DR   P-PSDB; AAE13823.
XX
PT   New human lung-specific polynucleotides and polypeptides for the
PT   diagnosis and treatment of disease e.g. lung cancer.
XX
PS   Example 4; Page 212; 378pp; English.
XX
CC   The invention relates to isolated lung tumour-specific proteins and their
CC   corresponding cDNA molecules. Lung tumour-specific proteins and their
CC   antigen-presenting cells are useful for stimulating and/or expanding T
CC   cells specific for a tumour protein, and for inhibiting the development
CC   of cancer. The invention also relates to a composition useful for
CC   stimulating an immune response, and for treating cancer. The lung tumour
CC   specific oligonucleotide is useful in gene therapy and for diagnosis,
CC   detection and treatment of lung cancer. The present sequence is a cDNA
CC   encoding human lung tumour-specific protein
XX
SQ   Sequence 547 BP; 144 A; 145 C; 146 G; 112 T; 0 U; 0 Other;

  Query Match           20.5%;  Score 449;  DB 4;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 1.9e-212;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108

Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168

Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228

Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288

Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348

Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408

Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468

Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528

Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548

Db        528 AGATCTTCGTGAAGACCCTG 547


RESULT 25
ADD66479
ID   ADD66479 standard; cDNA; 547 BP.
XX
AC   ADD66479;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human lung tumour-specific related cDNA, SEQ ID No 171.
```

XX
KW expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
KW human; lung tumour-specific; ss.
XX
OS Homo sapiens.
XX
PN WO200292001-A2.
XX
PD 21-NOV-2002.
XX
PF 10-MAY-2002; 2002WO-US014975.
XX
PR 11-MAY-2001; 2001US-00854133.
XX
PA (CORI-) CORIXA CORP.
XX
PI Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
XX
DR WPI; 2003-120592/11.
XX
PT New polynucleotide and polypeptide, useful for preparing a composition
PT for diagnosing, treating or preventing cancer.
XX
PS Example 4; SEQ ID NO 171; 494pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising one
CC of 32 47-6080 base pair sequences, given in the specification, or their
CC complements or degenerate variants, at least 20 contiguous residues of a
CC sequence in, or having at least 75 or 90 % identity with the isolated
CC polynucleotide, or that hybridise with the polynucleotide. The invention
CC further comprises: an isolated polypeptide; an expression vector
CC comprising the polynucleotide operably linked to an expression control
CC sequence; a host cell transformed or transfected with the expression
CC vector; an isolated antibody or its antigen-binding fragment that
CC specifically binds to the polypeptide; a method for detecting the
CC presence of a cancer in a patient; a fusion protein comprising the
CC polypeptide; an oligonucleotide that hybridises to the isolated
CC polynucleotide under moderately stringent conditions; a method for
CC stimulating and/or expanding T cells specific for a tumour protein; an
CC isolated T cell population; a composition comprising a first component
CC consisting of carriers and immunostimulants and a second component; a
CC method for stimulating an immune response in a patient; a method for
CC treating cancer in a patient; a method for determining cancer in a
CC patient; a diagnostic kit comprising at least one oligonucleotide or
CC antibody and a detection reagent comprising a reporter group; and a
CC method for inhibiting the development of cancer in a patient. The
CC compositions of the invention have cytostatic activity and can be used to
CC create a vaccine. The isolated polynucleotide is useful for preparing a
CC composition for diagnosing, treating or preventing cancer. This
CC polynucleotide sequence represents a human lung tumour-specific cDNA
CC sequence relating to the invention.
XX
SQ Sequence 547 BP; 144 A; 145 C; 146 G; 112 T; 0 U; 0 Other;

Query Match 20.5%; Score 449; DB 10; Length 547;
Best Local Similarity 99.8%; Pred. No. 1.9e-212;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy         289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy         349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy         409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy         469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy         529 AGATCTTCGTGAAGACCCTG 548
               ||||||||||||||||||||
Db         528 AGATCTTCGTGAAGACCCTG 547
```

RESULT 26
ADE87733
ID ADE87733 standard; cDNA; 547 BP.
XX
AC ADE87733;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human lung tumour antigen cDNA #118.
XX
KW Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
KW T cell; immune response; immunostimulant; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003118599-A1.
XX
PD 26-JUN-2003.
XX
PF 10-MAY-2002; 2002US-00144649.
XX
PR 02-APR-1999; 99US-00285323.
PR 09-AUG-1999; 99US-00370838.
PR 30-DEC-1999; 99US-00476235.
PR 03-MAR-2000; 2000US-00518809.
PR 29-MAR-2000; 2000US-00538037.
PR 05-JUN-2000; 2000US-00588937.
PR 18-AUG-2000; 2000US-00640878.
PR 20-SEP-2000; 2000US-00667170.
PR 01-NOV-2000; 2000US-00704512.
PR 14-DEC-2000; 2000US-00738973.
PR 11-MAY-2001; 2001US-00854133.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Lodes MJ, Wang T, Fan L, Mcneill PD;
XX
DR WPI; 2003-897103/82.
DR P-PSDB; ADE87767.
XX
PT New polynucleotides encode lung tumor antigens and are useful to
PT stimulate an immune response or detect or treat a cancer in a patient,
PT particularly lung cancer.
XX
PS Example 4; SEQ ID NO 171; 63pp; English.
XX
CC The invention relates to polynucleotides encoding lung tumour antigens.
CC The invention also relates to the polypeptides encoded by the
CC polynucleotides, isolated antibodies or antigen-binding fragments that
CC specifically bind the polypeptides and a method for detecting cancer in a
CC patient, comprising obtaining a biological sample from the patient,

CC contacting the sample with a binding agent that binds a polypeptide of
CC the invention, detecting in the sample an amount of polypeptide that
CC binds to the binding agent, and comparing the amount of polypeptide to a
CC predetermined cut-off value. T cells specific for a tumour protein can be
CC stimulated and/or expanded by contacting the T cells with a polypeptide,
CC polynucleotide or an antigen-presenting cell that expresses a
CC polypeptide. Cancer development can be inhibited by incubating CD4+
CC and/or CD8+ T cells isolated from a patient with a polypeptide,
CC polynucleotide or an antigen-presenting cell that expresses a
CC polypeptide, so that the T cells proliferate. The invention is used to
CC stimulate an immune response or to detect or treat a cancer in a patient,
CC particularly lung cancer. This sequence represents human lung tumour
CC antigen cDNA of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification but was obtained in
CC electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 547 BP; 144 A; 145 C; 146 G; 112 T; 0 U; 0 Other;

Query Match 20.5%; Score 449; DB 10; Length 547;
Best Local Similarity 99.8%; Pred. No. 1.9e-212;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy       109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy       169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy       229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy       289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy       349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy       409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy       469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy       529 AGATCTTCGTGAAGACCCTG 548
             ||||||||||||||||||||
Db       528 AGATCTTCGTGAAGACCCTG 547
```

RESULT 27
AAA77883
ID AAA77883 standard; cDNA; 460 BP.
XX
AC AAA77883;
XX
DT 14-NOV-2000 (first entry)
XX
DE cDNA encoding human colon tumour polypeptide, SEQ ID NO:163.
XX
KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
KW immunotherapy; diagnosis; progression; ss.
XX
OS Homo sapiens.
XX
PN WO200037643-A2.
XX
PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US030909.
XX
PR 23-DEC-1998; 98US-00221298.
PR 02-JUL-1999; 99US-00347496.
PR 22-SEP-1999; 99US-00401064.
PR 19-NOV-1999; 99US-00444242.
PR 02-DEC-1999; 99US-00454150.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk J;
PI Wang T, Yuqiu J;
XX
DR WPI; 2000-442671/38.
XX
PT New colon tumor polypeptides used to inhibit the development of cancer,
PT especially colon cancer, and for diagnosing and monitoring the
PT progression of the cancer.
XX
PS Claim 29; Page 138; 229pp; English.
XX
CC Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
CC portions of proteins which are associated with human colon tumours. The
CC invention also specifically discloses 8 human colon tumour proteins
CC (AAB11897-B11904). The nucleic acids, the polypeptides they encode, and
CC antigen presenting cells (APCs, preferably dendritic cells) expressing
CC such polypeptides may be used in vaccines that target tumour cells,
CC especially colon tumour cells, thereby inhibiting the development of
CC cancer. T-cells specific for the polypeptide expressed by the APC are
CC used to remove tumour cells from biological samples, especially blood or
CC fractions thereof. The sample or the isolated T-cells specific for the
CC polypeptide can then be used to inhibit cancer development. CD4+ and/or
CC CD8+ T-cells from a patient may be incubated with a polypeptide or
CC nucleic acid of the invention, or an APC expressing such a polypeptide,
CC to cause the proliferation of specific T-cells. The T-cells can be cloned
CC and then administered back to the patient to inhibit cancer development.
CC Nucleic acids encoding the polypeptides and antibodies against the
CC polypeptides may be used to determine the expression level of a tumour
CC protein of the invention, and therefore to determine whether cancer cells
CC are present. Such diagnostic methods may also be used to monitor the
CC progression of a cancer by repeating the processes at time intervals, and
CC comparing the current result to previous results. The present sequence
CC represents a cDNA encoding a human colon tumour polypeptide
XX
SQ Sequence 460 BP; 126 A; 115 C; 123 G; 96 T; 0 U; 0 Other;

Query Match 19.8%; Score 433; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.7e-204;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy        83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy       143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy       203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy       263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 28
AAI28621
ID   AAI28621 standard; cDNA; 460 BP.
XX
AC   AAI28621;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Colon tumour related determined cDNA sequence for CT22.
XX
KW   Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
KW   gene therapy; vaccine; colonic cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200149716-A2.
XX
PD   12-JUL-2001.
XX
PF   29-DEC-2000; 2000WO-US035596.
XX
PR   30-DEC-1999;   99US-00476296.
PR   10-JAN-2000; 2000US-00480321.
PR   15-FEB-2000; 2000US-00504629.
PR   06-MAR-2000; 2000US-00519444.
PR   19-MAY-2000; 2000US-00575251.
PR   29-JUN-2000; 2000US-00609448.
PR   28-AUG-2000; 2000US-00649811.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Lodes MJ,  Secrist H,  Benson DR,  Meagher MJ,  Stolk JA;
PI   King GE,  Wang T,  Jiang Y;
XX
DR   WPI; 2001-441847/47.
XX
PT   Colon tumor associated proteins and nucleic acids useful for the
PT   prevention, diagnosis and treatment of colonic cancer.
XX
PS   Claim 25; Page 179; 472pp; English.
XX
CC   The present invention describes colon tumour associated proteins (I) and
CC   the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC   (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC   (II) may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate colon tumour associated protein (TCAP)
CC   expression, such as colonic cancer. For example, (I) and (II) may be used
CC   to treat disorders associated with decreased expression by rectifying
CC   mutations or deletions in a patient's genome that affect the activity of
CC   TCAPs by expressing inactive proteins or to supplement the patients own
CC   production of them. Additionally, (II) may be used to produce the TCAP
CC   proteins, by inserting the nucleic acids into a host cell culturing the
CC   cell to express the protein. (II) and its complementary sequences may
CC   also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
CC   and hybridisation assays to detect and quantitate the presence of similar
CC   nucleic acids in samples, and therefore which patients may be in need of
CC   restorative therapy. (I) may also be used as antigens in the production
CC   of antibodies against TCAPs and in assays to identify modulators of TCAP
CC   expression and activity. Anti-(I) antibodies and antagonists may also be
CC   used to down regulate TCAP expression and activity. The anti-(I)
CC   antibodies may also be used as diagnostic agents for detecting the
CC   presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay
CC   (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent
CC   nucleotide and amino acid sequences given in the exemplification of the
CC   present invention
XX
SQ   Sequence 460 BP; 126 A; 115 C; 123 G; 96 T; 0 U; 0 Other;

  Query Match             19.8%;  Score 433;  DB 4;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-204;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 29
ABZ32807
ID   ABZ32807 standard; cDNA; 460 BP.
XX
AC   ABZ32807;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Human colon tumour related cDNA for CT22 SEQ ID NO:163.
XX
KW   Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
KW   tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200283070-A2.
XX
PD   24-OCT-2002.
XX
PF   09-APR-2002; 2002WO-US011475.
XX
PR   10-APR-2001; 2001US-00833263.
PR   03-AUG-2001; 2001US-00922217.
PR   19-DEC-2001; 2001US-00025380.
XX
PA   (CORI-) CORIXA CORP.
```

```
XX
PI   Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI   Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD, Skeiky YAW;
PI   Fanger GR, Vedvick TS, Carter D;
XX
DR   WPI; 2003-067548/06.
XX
PT   New polynucleotide, useful for the preparation of a composition for
PT   stimulating an immune response against, or treating, cancer.
XX
PS   Example 3; Page 181; 537pp; English.
XX
CC   The present invention describes compounds (I) for the immunotherapy and
CC   diagnosis of colon cancer. Also described: (1) a method for detecting the
CC   presence of cancer in a patient; (2) a method for stimulating and/or
CC   expanding T cells specific for a tumour protein; (3) an isolated T cell
CC   population comprising T cells prepared by the method of (2); (4) a method
CC   for stimulating an immune response in a patient; (5) a method for
CC   treating cancer in a patient; and (6) a method for inhibiting the
CC   development of cancer in a patient. (I) have immunostimulant and
CC   cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
CC   and ABP55343 to ABP55391 represent human colon cancer/tumour related
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 460 BP; 126 A; 115 C; 123 G; 96 T; 0 U; 0 Other;

  Query Match             19.8%;  Score 433;  DB 8;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-204;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        -88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 30
AAC79233
ID   AAC79233 standard; cDNA; 494 BP.
XX
AC   AAC79233;
XX
DT   05-FEB-2001  (first entry)
XX
DE   Human lung tumour-specific cDNA #186.
XX
KW   Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200060077-A2.
XX
PD   12-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US008560.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG, Lodes MJ, Mohamath R, Secrist H;
XX
DR   WPI; 2000-638466/61.
XX
PT   Novel lung tumor polypeptides and polynucleotides, useful for detecting,
PT   monitoring or treating cancer, especially lung cancer.
XX
PS   Claim 24; Page 205; 243pp; English.
XX
CC   The present sequence is given in a specification relating to compounds
CC   for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC   least an immunogenic part of a lung tumour protein are disclosed. The
CC   polypeptides are useful for inhibiting the development of cancer,
CC   especially lung cancer. Samples of T cells expressing the polypeptides
CC   may be used to inhibit the development of cancer. The polypeptides are
CC   also useful for detecting and monitoring the progression of cancer,
CC   especially lung cancer
XX
SQ   Sequence 494 BP; 134 A; 133 C; 129 G; 97 T; 0 U; 1 Other;

  Query Match             19.6%;  Score 430;  DB 3;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 5.2e-203;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493
```

```
Qy        504 G 504
              |
Db        494 G 494
```

RESULT 31
AAD23309
ID AAD23309 standard; cDNA; 494 BP.
XX
AC AAD23309;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human lung tumour-specific 2LT-118 cDNA.
XX
KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 2LT-118; ss.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US009991.
XX
PR 29-MAR-2000; 2000US-00538037.
PR 05-JUN-2000; 2000US-00588937.
PR 18-AUG-2000; 2000US-00640878.
PR 22-SEP-2000; 2000US-0234517P.
PR 01-NOV-2000; 2000US-00704512.
PR 14-DEC-2000; 2000US-00738973.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
DR WPI; 2001-639201/73.
XX
PT New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer.
XX
PS Example 4; Page 256; 378pp; English.
XX
CC The invention relates to isolated lung tumour-specific proteins and their
CC corresponding cDNA molecules. Lung tumour-specific proteins and their
CC antigen-presenting cells are useful for stimulating and/or expanding T
CC cells specific for a tumour protein, and for inhibiting the development
CC of cancer. The invention also relates to a composition useful for
CC stimulating an immune response, and for treating cancer. The lung tumour
CC specific oligonucleotide is useful in gene therapy and for diagnosis,
CC detection and treatment of lung cancer. The present sequence is human
CC lung tumour-specific cDNA
XX
SQ Sequence 494 BP; 134 A; 133 C; 129 G; 97 T; 0 U; 1 Other;

  Query Match   19.6%; Score 430; DB 4; Length 494;
  Best Local Similarity   99.8%; Pred. No. 5.2e-203;
  Matches 480; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

```
Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494
```

RESULT 32
ADD66583
ID ADD66583 standard; cDNA; 494 BP.
XX
AC ADD66583;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human lung tumour-specific related cDNA, SEQ ID No 275.
XX
KW expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
KW human; lung tumour-specific; ss.
XX
OS Homo sapiens.
XX
PN WO200292001-A2.
XX
PD 21-NOV-2002.
XX
PF 10-MAY-2002; 2002WO-US014975.
XX
PR 11-MAY-2001; 2001US-00854133.
XX
PA (CORI-) CORIXA CORP.
XX
PI Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
XX
DR WPI; 2003-120592/11.
XX
PT New polynucleotide and polypeptide, useful for preparing a composition
PT for diagnosing, treating or preventing cancer.
XX
PS Example 4; SEQ ID NO 275; 494pp; English.
XX
CC The invention relates to a novel isolated polynucleotide comprising one
CC of 32 47-6080 base pair sequences, given in the specification, or their
CC complements or degenerate variants, at least 20 contiguous residues of a
CC sequence in, or having at least 75 or 90 % identity with the isolated
CC polynucleotide, or that hybridise with the polynucleotide. The invention
CC further comprises: an isolated polypeptide; an expression vector
CC comprising the polynucleotide operably linked to an expression control
CC sequence; a host cell transformed or transfected with the expression
CC vector; an isolated antibody or its antigen-binding fragment that
CC specifically binds to the polypeptide; a method for detecting the
CC presence of a cancer in a patient; a fusion protein comprising the
CC polypeptide; an oligonucleotide that hybridises to the isolated
CC polynucleotide under moderately stringent conditions; a method for
CC stimulating and/or expanding T cells specific for a tumour protein; an

```
CC   isolated T cell population; a composition comprising a first component
CC   consisting of carriers and immunostimulants and a second component; a
CC   method for stimulating an immune response in a patient; a method for
CC   treating cancer in a patient; a method for determining cancer in a
CC   patient; a diagnostic kit comprising at least one oligonucleotide or
CC   antibody and a detection reagent comprising a reporter group; and a
CC   method for inhibiting the development of cancer in a patient. The
CC   compositions of the invention have cytostatic activity and can be used to
CC   create a vaccine. The isolated polynucleotide is useful for preparing a
CC   composition for diagnosing, treating or preventing cancer. This
CC   polynucleotide sequence represents a human lung tumour-specific cDNA
CC   sequence relating to the invention.
XX
SQ   Sequence 494 BP; 134 A; 133 C; 129 G; 97 T; 0 U; 1 Other;

  Query Match             19.6%;  Score 430;  DB 10;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 5.2e-203;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494


RESULT 33
ADE87837
ID   ADE87837 standard; cDNA; 494 BP.
XX
AC   ADE87837;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human lung tumour antigen cDNA #187.
XX
KW   Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
KW   T cell; immune response; immunostimulant; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2003118599-A1.
XX
PD   26-JUN-2003.
XX
PF   10-MAY-2002; 2002US-00144649.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   20-SEP-2000; 2000US-00667170.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Lodes MJ,  Wang T,  Fan L,  Mcneill PD;
XX
DR   WPI; 2003-897103/82.
XX
PT   New polynucleotides encode lung tumor antigens and are useful to
PT   stimulate an immune response or detect or treat a cancer in a patient,
PT   particularly lung cancer.
XX
PS   Example 4; SEQ ID NO 275; 63pp; English.
XX
CC   The invention relates to polynucleotides encoding lung tumour antigens.
CC   The invention also relates to the polypeptides encoded by the
CC   polynucleotides, isolated antibodies or antigen-binding fragments that
CC   specifically bind the polypeptides and a method for detecting cancer in a
CC   patient, comprising obtaining a biological sample from the patient,
CC   contacting the sample with a binding agent that binds a polypeptide of
CC   the invention, detecting in the sample an amount of polypeptide that
CC   binds to the binding agent, and comparing the amount of polypeptide to a
CC   predetermined cut-off value. T cells specific for a tumour protein can be
CC   stimulated and/or expanded by contacting the T cells with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide. Cancer development can be inhibited by incubating CD4+
CC   and/or CD8+ T cells isolated from a patient with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide, so that the T cells proliferate. The invention is used to
CC   stimulate an immune response or to detect or treat a cancer in a patient,
CC   particularly lung cancer. This sequence represents human lung tumour
CC   antigen cDNA of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification but was obtained in
CC   electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 494 BP; 134 A; 133 C; 129 G; 97 T; 0 U; 1 Other;

  Query Match             19.6%;  Score 430;  DB 10;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 5.2e-203;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494
```

RESULT 34
AAZ07244
ID AAZ07244 standard; cDNA; 483 BP.
XX
AC AAZ07244;
XX
DT 13-OCT-1999 (first entry)
XX
DE Human lung tumour protein PSLT-041 cDNA sequence.
XX
KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW immunotherapy; detection; inhibition; ss.
XX
OS Homo sapiens.
XX
PN WO9938973-A2.
XX
PD 05-AUG-1999.
XX
PF 26-JAN-1999; 99WO-US001642.
XX
PR 28-JAN-1998; 98US-00015022.
PR 28-JAN-1998; 98US-00015029.
PR 18-MAR-1998; 98US-00040828.
PR 18-MAR-1998; 98US-00040831.
PR 23-JUL-1998; 98US-00122191.
PR 23-JUL-1998; 98US-00122192.
PR 22-DEC-1998; 98US-00219245.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Frudakis TN, Mohamath R;
XX
DR WPI; 1999-479187/40.
XX
PT Lung tumor specific polynucleotides for inhibiting the development of
PT lung cancer.
XX
PS Claim 13; Page 130; 171pp; English.
XX
CC The present invention describes lung tumour specific polynucleotides and
CC tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325 represent
CC specifically claimed polynucleotides, and AAY29486 to AAY29571 represent
CC amino acid sequences from the present invention. The lung tumour specific
CC polynucleotides and polypeptides can be used in pharmaceutical
CC compositions and vaccines to inhibit the development of lung cancer. They
CC can also be used to detect lung cancer in a patient. Probes and
CC antibodies derived from the lung tumour sequences are useful in detection
CC of lung cancer
XX
SQ Sequence 483 BP; 135 A; 128 C; 131 G; 89 T; 0 U; 0 Other;

```
  Query Match            19.4%;  Score 426;  DB 2;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 5e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483
```

RESULT 35
AAC79135
ID AAC79135 standard; cDNA; 483 BP.
XX
AC AAC79135;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human lung tumour-specific cDNA #88.
XX
KW Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200060077-A2.
XX
PD 12-OCT-2000.
XX
PF 30-MAR-2000; 2000WO-US008560.
XX
PR 02-APR-1999; 99US-00285323.
PR 09-AUG-1999; 99US-00370838.
PR 30-DEC-1999; 99US-00476235.
PR 03-MAR-2000; 2000US-00518809.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H;
XX
DR WPI; 2000-638466/61.
XX
PT Novel lung tumor polypeptides and polynucleotides, useful for detecting,
PT monitoring or treating cancer, especially lung cancer.
XX
PS Claim 1; Page 145-146; 243pp; English.
XX
CC The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at

```
CC   least an immunogenic part of a lung tumour protein are disclosed. The
CC   polypeptides are useful for inhibiting the development of cancer,
CC   especially lung cancer. Samples of T cells expressing the polypeptides
CC   may be used to inhibit the development of cancer. The polypeptides are
CC   also useful for detecting and monitoring the progression of cancer,
CC   especially lung cancer
XX
SQ   Sequence 483 BP; 135 A; 128 C; 131 G; 89 T; 0 U; 0 Other;

  Query Match           19.4%;  Score 426;  DB 3;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 5e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 36
AAD23210
ID   AAD23210 standard; cDNA; 483 BP.
XX
AC   AAD23210;
XX
DT   26-FEB-2002  (first entry)
XX
DE   Human lung tumour-specific PSLT-041 cDNA.
XX
KW   Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW   antisense-therapy; vaccine; immune response; lung cancer; PSLT-041; ss.
XX
OS   Homo sapiens.
XX
PN   WO200172295-A2.
XX
PD   04-OCT-2001.
XX
PF   28-MAR-2001; 2001WO-US009991.
XX
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   22-SEP-2000; 2000US-0234517P.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI   Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
DR   WPI; 2001-639201/73.
XX
PT   New human lung-specific polynucleotides and polypeptides for the
PT   diagnosis and treatment of disease e.g. lung cancer.
XX
PS   Example 3; Page 199-200; 378pp; English.
XX
CC   The invention relates to isolated lung tumour-specific proteins and their
CC   corresponding cDNA molecules. Lung tumour-specific proteins and their
CC   antigen-presenting cells are useful for stimulating and/or expanding T
CC   cells specific for a tumour protein, and for inhibiting the development
CC   of cancer. The invention also relates to a composition useful for
CC   stimulating an immune response, and for treating cancer. The lung tumour
CC   specific oligonucleotide is useful in gene therapy and for diagnosis,
CC   detection and treatment of lung cancer. The present sequence is human
CC   lung tumour-specific cDNA
XX
SQ   Sequence 483 BP; 135 A; 128 C; 131 G; 89 T; 0 U; 0 Other;

  Query Match           19.4%;  Score 426;  DB 4;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 5e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 37
ADD66449
ID   ADD66449 standard; cDNA; 483 BP.
XX
AC   ADD66449;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human lung tumour-specific related cDNA, SEQ ID No 141.
XX
KW   expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
```

```
KW   human; lung tumour-specific; ss.
XX
OS   Homo sapiens.
XX
PN   WO200292001-A2.
XX
PD   21-NOV-2002.
XX
PF   10-MAY-2002; 2002WO-US014975.
XX
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Lodes MJ,  Wang T,  Fan L,  Algate PA,  Mcneill PD;
XX
DR   WPI; 2003-120592/11.
XX
PT   New polynucleotide and polypeptide, useful for preparing a composition
PT   for diagnosing, treating or preventing cancer.
XX
PS   Example 3; SEQ ID NO 141; 494pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising one
CC   of 32 47-6080 base pair sequences, given in the specification, or their
CC   complements or degenerate variants, at least 20 contiguous residues of a
CC   sequence in, or having at least 75 or 90 % identity with the isolated
CC   polynucleotide, or that hybridise with the polynucleotide. The invention
CC   further comprises: an isolated polypeptide; an expression vector
CC   comprising the polynucleotide operably linked to an expression control
CC   sequence; a host cell transformed or transfected with the expression
CC   vector; an isolated antibody or its antigen-binding fragment that
CC   specifically binds to the polypeptide; a method for detecting the
CC   presence of a cancer in a patient; a fusion protein comprising the
CC   polypeptide; an oligonucleotide that hybridises to the isolated
CC   polynucleotide under moderately stringent conditions; a method for
CC   stimulating and/or expanding T cells specific for a tumour protein; an
CC   isolated T cell population; a composition comprising a first component
CC   consisting of carriers and immunostimulants and a second component; a
CC   method for stimulating an immune response in a patient; a method for
CC   treating cancer in a patient; a method for determining cancer in a
CC   patient; a diagnostic kit comprising at least one oligonucleotide or
CC   antibody and a detection reagent comprising a reporter group; and a
CC   method for inhibiting the development of cancer in a patient. The
CC   compositions of the invention have cytostatic activity and can be used to
CC   create a vaccine. The isolated polynucleotide is useful for preparing a
CC   composition for diagnosing, treating or preventing cancer. This
CC   polynucleotide sequence represents a human lung tumour-specific cDNA
CC   sequence relating to the invention.
XX
SQ   Sequence 483 BP; 135 A; 128 C; 131 G; 89 T; 0 U; 0 Other;

  Query Match            19.4%;  Score 426;  DB 10;  Length 483;
  Best Local Similarity  100.0%;  Pred. No. 5e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 38
ADE87703
ID   ADE87703 standard; cDNA; 483 BP.
XX
AC   ADE87703;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human lung tumour antigen cDNA #88.
XX
KW   Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
KW   T cell; immune response; immunostimulant; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2003118599-A1.
XX
PD   26-JUN-2003.
XX
PF   10-MAY-2002; 2002US-00144649.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   20-SEP-2000; 2000US-00667170.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Lodes MJ,  Wang T,  Fan L,  Mcneill PD;
XX
DR   WPI; 2003-897103/82.
XX
PT   New polynucleotides encode lung tumor antigens and are useful to
PT   stimulate an immune response or detect or treat a cancer in a patient,
PT   particularly lung cancer.
XX
PS   Example 3; SEQ ID NO 141; 63pp; English.
XX
CC   The invention relates to polynucleotides encoding lung tumour antigens.
CC   The invention also relates to the polypeptides encoded by the
CC   polynucleotides, isolated antibodies or antigen-binding fragments that
CC   specifically bind the polypeptides and a method for detecting cancer in a
CC   patient, comprising obtaining a biological sample from the patient,
CC   contacting the sample with a binding agent that binds a polypeptide of
CC   the invention, detecting in the sample an amount of polypeptide that
CC   binds to the binding agent, and comparing the amount of polypeptide to a
CC   predetermined cut-off value. T cells specific for a tumour protein can be
CC   stimulated and/or expanded by contacting the T cells with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide. Cancer development can be inhibited by incubating CD4+
```

```
CC   and/or CD8+ T cells isolated from a patient with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide, so that the T cells proliferate. The invention is used to
CC   stimulate an immune response or to detect or treat a cancer in a patient,
CC   particularly lung cancer. This sequence represents human lung tumour
CC   antigen cDNA of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification but was obtained in
CC   electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 483 BP; 135 A; 128 C; 131 G; 89 T; 0 U; 0 Other;

  Query Match             19.4%;  Score 426;  DB 10;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 5e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 39
ADD67000
ID   ADD67000 standard; cDNA; 412 BP.
XX
AC   ADD67000;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human lung tumour-specific related cDNA, SEQ ID No 692.
XX
KW   expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
KW   human; lung tumour-specific; ss.
XX
OS   Homo sapiens.
XX
PN   WO200292001-A2.
XX
PD   21-NOV-2002.
XX
PF   10-MAY-2002; 2002WO-US014975.
XX
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Lodes MJ,  Wang T,  Fan L,  Algate PA,  Mcneill PD;
XX
DR   WPI; 2003-120592/11.
XX
PT   New polynucleotide and polypeptide, useful for preparing a composition
PT   for diagnosing, treating or preventing cancer.
XX
PS   Claim 11; SEQ ID NO 692; 494pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising one
CC   of 32 47-6080 base pair sequences, given in the specification, or their
CC   complements or degenerate variants, at least 20 contiguous residues of a
CC   sequence in, or having at least 75 or 90 % identity with the isolated
CC   polynucleotide, or that hybridise with the polynucleotide. The invention
CC   further comprises: an isolated polypeptide; an expression vector
CC   comprising the polynucleotide operably linked to an expression control
CC   sequence; a host cell transformed or transfected with the expression
CC   vector; an isolated antibody or its antigen-binding fragment that
CC   specifically binds to the polypeptide; a method for detecting the
CC   presence of a cancer in a patient; a fusion protein comprising the
CC   polypeptide; an oligonucleotide that hybridises to the isolated
CC   polynucleotide under moderately stringent conditions; a method for
CC   stimulating and/or expanding T cells specific for a tumour protein; an
CC   isolated T cell population; a composition comprising a first component
CC   consisting of carriers and immunostimulants and a second component; a
CC   method for stimulating an immune response in a patient; a method for
CC   treating cancer in a patient; a method for determining cancer in a
CC   patient; a diagnostic kit comprising at least one oligonucleotide or
CC   antibody and a detection reagent comprising a reporter group; and a
CC   method for inhibiting the development of cancer in a patient. The
CC   compositions of the invention have cytostatic activity and can be used to
CC   create a vaccine. The isolated polynucleotide is useful for preparing a
CC   composition for diagnosing, treating or preventing cancer. This
CC   polynucleotide sequence represents a human lung tumour-specific cDNA
CC   sequence relating to the invention.
XX
SQ   Sequence 412 BP; 114 A; 109 C; 107 G; 82 T; 0 U; 0 Other;

  Query Match             18.8%;  Score 411;  DB 10;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-193;
  Matches  411;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         35 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 61

Qy         95 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 121

Qy        155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 181

Qy        215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 241

Qy        275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 301

Qy        335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 361

Qy        395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 412


RESULT 40
ADE88254
```

```
ID   ADE88254 standard; cDNA; 412 BP.
XX
AC   ADE88254;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human lung tumour antigen cDNA #587.
XX
KW   Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
KW   T cell; immune response; immunostimulant; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2003118599-A1.
XX
PD   26-JUN-2003.
XX
PF   10-MAY-2002; 2002US-00144649.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   20-SEP-2000; 2000US-00667170.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Lodes MJ,  Wang T,  Fan L,  Mcneill PD;
XX
DR   WPI; 2003-897103/82.
XX
PT   New polynucleotides encode lung tumor antigens and are useful to
PT   stimulate an immune response or detect or treat a cancer in a patient,
PT   particularly lung cancer.
XX
PS   Claim 11; SEQ ID NO 692; 63pp; English.
XX
CC   The invention relates to polynucleotides encoding lung tumour antigens.
CC   The invention also relates to the polypeptides encoded by the
CC   polynucleotides, isolated antibodies or antigen-binding fragments that
CC   specifically bind the polypeptides and a method for detecting cancer in a
CC   patient, comprising obtaining a biological sample from the patient,
CC   contacting the sample with a binding agent that binds a polypeptide of
CC   the invention, detecting in the sample an amount of polypeptide that
CC   binds to the binding agent, and comparing the amount of polypeptide to a
CC   predetermined cut-off value. T cells specific for a tumour protein can be
CC   stimulated and/or expanded by contacting the T cells with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide. Cancer development can be inhibited by incubating CD4+
CC   and/or CD8+ T cells isolated from a patient with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide, so that the T cells proliferate. The invention is used to
CC   stimulate an immune response or to detect or treat a cancer in a patient,
CC   particularly lung cancer. This sequence represents human lung tumour
CC   antigen cDNA of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification but was obtained in
CC   electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 412 BP; 114 A; 109 C; 107 G; 82 T; 0 U; 0 Other;

  Query Match             18.8%;  Score 411;  DB 10;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-193;
  Matches  411;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         35 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 61

Qy         95 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 121

Qy        155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 181

Qy        215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 241

Qy        275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 301

Qy        335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 361

Qy        395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 412


RESULT 41
AAC79206
ID   AAC79206 standard; cDNA; 434 BP.
XX
AC   AAC79206;
XX
DT   05-FEB-2001  (first entry)
XX
DE   Human lung tumour-specific cDNA #159.
XX
KW   Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200060077-A2.
XX
PD   12-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US008560.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Lodes MJ,  Mohamath R,  Secrist H;
XX
DR   WPI; 2000-638466/61.
XX
PT   Novel lung tumor polypeptides and polynucleotides, useful for detecting,
PT   monitoring or treating cancer, especially lung cancer.
XX
PS   Claim 1; Page 196-197; 243pp; English.
XX
CC   The present sequence is given in a specification relating to compounds
CC   for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC   least an immunogenic part of a lung tumour protein are disclosed. The
CC   polypeptides are useful for inhibiting the development of cancer,
CC   especially lung cancer. Samples of T cells expressing the polypeptides
CC   may be used to inhibit the development of cancer. The polypeptides are
CC   also useful for detecting and monitoring the progression of cancer,
CC   especially lung cancer
XX
SQ   Sequence 434 BP; 118 A; 115 C; 116 G; 85 T; 0 U; 0 Other;
```

```
  Query Match             18.7%;  Score 410;  DB 3;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-193;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410


RESULT 42
AAD23282
ID   AAD23282 standard; cDNA; 434 BP.
XX
AC   AAD23282;
XX
DT   26-FEB-2002  (first entry)
XX
DE   Human lung tumour-specific 2LT-55 cDNA.
XX
KW   Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW   antisense-therapy; vaccine; immune response; lung cancer; 2LT-55; ss.
XX
OS   Homo sapiens.
XX
PN   WO200172295-A2.
XX
PD   04-OCT-2001.
XX
PF   28-MAR-2001; 2001WO-US009991.
XX
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   22-SEP-2000; 2000US-0234517P.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Lodes MJ,  Mohamath R,  Secrist H,  Benson DR,  Indirias CY;
PI   Henderson RA,  Fling SP,  Algate PA,  Elliot M,  Mannion J,  Kalos MD;
XX
DR   WPI; 2001-639201/73.
XX
PT   New human lung-specific polynucleotides and polypeptides for the
PT   diagnosis and treatment of disease e.g. lung cancer.
XX
PS   Example 4; Page 248; 378pp; English.
XX
CC   The invention relates to isolated lung tumour-specific proteins and their
CC   corresponding cDNA molecules. Lung tumour-specific proteins and their
CC   antigen-presenting cells are useful for stimulating and/or expanding T
CC   cells specific for a tumour protein, and for inhibiting the development
CC   of cancer. The invention also relates to a composition useful for
CC   stimulating an immune response, and for treating cancer. The lung tumour
CC   specific oligonucleotide is useful in gene therapy and for diagnosis,
CC   detection and treatment of lung cancer. The present sequence is human
CC   lung tumour-specific cDNA
XX
SQ   Sequence 434 BP; 118 A; 115 C; 116 G; 85 T; 0 U; 0 Other;

  Query Match             18.7%;  Score 410;  DB 4;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-193;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410


RESULT 43
ADD66556
ID   ADD66556 standard; cDNA; 434 BP.
XX
AC   ADD66556;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human lung tumour-specific related cDNA, SEQ ID No 248.
XX
KW   expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
KW   human; lung tumour-specific; ss.
XX
OS   Homo sapiens.
XX
PN   WO200292001-A2.
XX
PD   21-NOV-2002.
XX
PF   10-MAY-2002; 2002WO-US014975.
XX
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Lodes MJ,  Wang T,  Fan L,  Algate PA,  Mcneill PD;
XX
DR   WPI; 2003-120592/11.
```

```
XX
PT   New polynucleotide and polypeptide, useful for preparing a composition
PT   for diagnosing, treating or preventing cancer.
XX
PS   Example 4; SEQ ID NO 248; 494pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide comprising one
CC   of 32 47-6080 base pair sequences, given in the specification, or their
CC   complements or degenerate variants, at least 20 contiguous residues of a
CC   sequence in, or having at least 75 or 90 % identity with the isolated
CC   polynucleotide, or that hybridise with the polynucleotide. The invention
CC   further comprises: an isolated polypeptide; an expression vector
CC   comprising the polynucleotide operably linked to an expression control
CC   sequence; a host cell transformed or transfected with the expression
CC   vector; an isolated antibody or its antigen-binding fragment that
CC   specifically binds to the polypeptide; a method for detecting the
CC   presence of a cancer in a patient; a fusion protein comprising the
CC   polypeptide; an oligonucleotide that hybridises to the isolated
CC   polynucleotide under moderately stringent conditions; a method for
CC   stimulating and/or expanding T cells specific for a tumour protein; an
CC   isolated T cell population; a composition comprising a first component
CC   consisting of carriers and immunostimulants and a second component; a
CC   method for stimulating an immune response in a patient; a method for
CC   treating cancer in a patient; a method for determining cancer in a
CC   patient; a diagnostic kit comprising at least one oligonucleotide or
CC   antibody and a detection reagent comprising a reporter group; and a
CC   method for inhibiting the development of cancer in a patient. The
CC   compositions of the invention have cytostatic activity and can be used to
CC   create a vaccine. The isolated polynucleotide is useful for preparing a
CC   composition for diagnosing, treating or preventing cancer. This
CC   polynucleotide sequence represents a human lung tumour-specific cDNA
CC   sequence relating to the invention.
XX
SQ   Sequence 434 BP; 118 A; 115 C; 116 G; 85 T; 0 U; 0 Other;

  Query Match             18.7%;  Score 410;  DB 10;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-193;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy          92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy         152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy         212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy         272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy         332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy         392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410


RESULT 44
ADE87810
ID   ADE87810 standard; cDNA; 434 BP.
XX
AC   ADE87810;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human lung tumour antigen cDNA #160.
XX
KW   Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
KW   T cell; immune response; immunostimulant; cytostatic.
XX
OS   Homo sapiens.
XX
PN   US2003118599-A1.
XX
PD   26-JUN-2003.
XX
PF   10-MAY-2002; 2002US-00144649.
XX
PR   02-APR-1999;   99US-00285323.
PR   09-AUG-1999;   99US-00370838.
PR   30-DEC-1999;   99US-00476235.
PR   03-MAR-2000; 2000US-00518809.
PR   29-MAR-2000; 2000US-00538037.
PR   05-JUN-2000; 2000US-00588937.
PR   18-AUG-2000; 2000US-00640878.
PR   20-SEP-2000; 2000US-00667170.
PR   01-NOV-2000; 2000US-00704512.
PR   14-DEC-2000; 2000US-00738973.
PR   11-MAY-2001; 2001US-00854133.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Lodes MJ,  Wang T,  Fan L,  Mcneill PD;
XX
DR   WPI; 2003-897103/82.
XX
PT   New polynucleotides encode lung tumor antigens and are useful to
PT   stimulate an immune response or detect or treat a cancer in a patient,
PT   particularly lung cancer.
XX
PS   Example 4; SEQ ID NO 248; 63pp; English.
XX
CC   The invention relates to polynucleotides encoding lung tumour antigens.
CC   The invention also relates to the polypeptides encoded by the
CC   polynucleotides, isolated antibodies or antigen-binding fragments that
CC   specifically bind the polypeptides and a method for detecting cancer in a
CC   patient, comprising obtaining a biological sample from the patient,
CC   contacting the sample with a binding agent that binds a polypeptide of
CC   the invention, detecting in the sample an amount of polypeptide that
CC   binds to the binding agent, and comparing the amount of polypeptide to a
CC   predetermined cut-off value. T cells specific for a tumour protein can be
CC   stimulated and/or expanded by contacting the T cells with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide. Cancer development can be inhibited by incubating CD4+
CC   and/or CD8+ T cells isolated from a patient with a polypeptide,
CC   polynucleotide or an antigen-presenting cell that expresses a
CC   polypeptide, so that the T cells proliferate. The invention is used to
CC   stimulate an immune response or to detect or treat a cancer in a patient,
CC   particularly lung cancer. This sequence represents human lung tumour
CC   antigen cDNA of the invention. Note: The sequence data for this patent
CC   did not form part of the printed specification but was obtained in
CC   electronic format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 434 BP; 118 A; 115 C; 116 G; 85 T; 0 U; 0 Other;

  Query Match             18.7%;  Score 410;  DB 10;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-193;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy          92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 45
ACH36460
ID ACH36460 standard; cDNA; 442 BP.
XX
AC ACH36460;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human endothelial cell cDNA #4593.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 23672; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 442 BP; 119 A; 114 C; 121 G; 88 T; 0 U; 0 Other;

```
  Query Match             18.6%;  Score 408;  DB 9;  Length 442;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-192;
  Matches  408;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1722 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        35 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 94

Qy      1782 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        95 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 154

Qy      1842 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       155 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 214

Qy      1902 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 1961
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       215 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 274

Qy      1962 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 2021
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       275 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 334

Qy      2022 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 2081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       335 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 394

Qy      2082 GAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 2129
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       395 GAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 442
```

RESULT 46
ACL55035/c
ID ACL55035 standard; cDNA; 743 BP.
XX
AC ACL55035;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:1170.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO2005000087-A2.
XX
PD 06-JAN-2005.
XX
PF 13-MAY-2004; 2004WO-US015421.
XX
PR 03-JUN-2003; 2003US-0475872P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
DR WPI; 2005-075421/08.
XX

PT New isolated polynucleotides, which are differentially expressed in colon
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT cancer, or pancreatic cancer.
XX
PS Claim 1; SEQ ID NO 1170; 97pp; English.
XX
CC The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC are differentially expressed in colon cancer cells. The invention also
CC relates to vectors and host cells comprising a differentially expressed
CC polynucleotide of the invention; a method for detecting a cancerous cell
CC by detection of a gene product of the polynucleotides; a method for
CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC of the polynucleotides; a method of treating an individual with cancer by
CC administration of a modulator of a gene product of the polynucleotides;
CC and an isolated antibody that specifically binds to a polypeptide encoded
CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC antibodies, and methods are useful for the detection of cancerous cells;
CC for the diagnosis, prognosis and management of cancer; for the
CC identification of agents that modulate the phenotype of cancerous cells;
CC for the identification of therapeutic targets for cancer chemotherapy;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 743 BP; 160 A; 190 C; 185 G; 207 T; 0 U; 1 Other;

Query Match 18.6%; Score 407; DB 14; Length 743;
Best Local Similarity 100.0%; Pred. No. 1.4e-191;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1353 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1412
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 615 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 556

Qy 1413 CTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1472
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 555 CTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 496

Qy 1473 ATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 1532
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 495 ATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 436

Qy 1533 AAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 1592
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 435 AAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 376

Qy 1593 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGT 1652
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 375 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGT 316

Qy 1653 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 1712
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 315 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 256

Qy 1713 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1759
|||||||||||||||||||||||||||||||||||||||||||||
Db 255 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 209

RESULT 47
ACN39602
ID ACN39602 standard; cDNA; 425 BP.
XX
AC ACN39602;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA325804, SEQ ID NO:3855.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 3855; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 425 BP; 118 A; 110 C; 112 G; 85 T; 0 U; 0 Other;

Query Match 18.4%; Score 403; DB 13; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.3e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1130 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 1189
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 60

Qy 1190 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1249
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 120

Qy 1250 TCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1309
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 TCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 180

```
Qy       1310 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 1369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 240

Qy       1370 CACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAG 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAG 300

Qy       1430 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGA 1489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGA 360

Qy       1490 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 1532
             |||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 403
```

RESULT 48
AAF21656
ID AAF21656 standard; DNA; 1187 BP.
XX
AC AAF21656;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human breast and ovarian cancer associated antigen gene SEQ ID 43.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS Homo sapiens.
XX
PN WO200055173-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005881.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-611515/58.
DR P-PSDB; AAB58753.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention, treatment
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
PT neurological diseases.
XX
PS Claim 1; Page 512; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive; nootropic;
CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial;
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC protein sequences are used in the diagnosis of cancer, particularly
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC and agonists may also be used in the diagnosis, prevention and treatment
CC of immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC cardiovascular disorders such as myocardial ischaemias; wound healing;
CC neurological diseases such as cerebral anoxia and epilepsy; and
CC infectious diseases
XX
SQ Sequence 1187 BP; 320 A; 302 C; 316 G; 232 T; 0 U; 17 Other;

```
  Query Match             18.2%;  Score 398;  DB 3;  Length 1187;
  Best Local Similarity   100.0%;  Pred. No. 4e-187;
  Matches  398;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 654

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 714

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1664
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 774

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 834

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 894

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 954

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1882
             |||||||||||||||||||||||||||||||||||||||
Db        955 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 992
```

RESULT 49
ACH19114
ID ACH19114 standard; cDNA; 389 BP.
XX
AC ACH19114;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult lung cDNA #117.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

```
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 6326; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 389 BP; 102 A; 102 C; 105 G; 80 T; 0 U; 0 Other;

  Query Match             17.2%;  Score 376;  DB 9;  Length 389;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-176;
  Matches  376;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1099 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACA 1158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACA 72

Qy       1159 ACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 1218
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 ACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 132

Qy       1219 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTG 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTG 192

Qy       1279 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGA 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGA 252

Qy       1339 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 1398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 312

Qy       1399 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCC 1458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCC 372

Qy       1459 TGACTGGTAAGACCAT 1474
             ||||||||||||||||
Db        373 TGACTGGTAAGACCAT 388


RESULT 50
ACH33450
ID   ACH33450 standard; cDNA; 506 BP.
XX
AC   ACH33450;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human endothelial cell cDNA #1583.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 20662; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 506 BP; 132 A; 127 C; 138 G; 106 T; 0 U; 3 Other;

  Query Match             17.2%;  Score 376;  DB 9;  Length 506;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-176;
  Matches  376;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1796 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1855
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         114 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 173

Qy       1856 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1915
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 233

Qy       1916 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 1975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         234 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 293

Qy       1976 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 2035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         294 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 353

Qy       2036 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 2095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         354 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 413
```

```
Qy       2096 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 2155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 473

Qy       2156 CATTGCACTTTCCTTT 2171
             ||||||||||||||||
Db        474 CATTGCACTTTCCTTT 489


RESULT 51
AAA69842
ID   AAA69842 standard; cDNA; 518 BP.
XX
AC   AAA69842;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Human ovarian carcinoma antigen polynucleotide SEQ ID NO:152.
XX
KW   Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW   tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200036107-A2.
XX
PD   22-JUN-2000.
XX
PF   17-DEC-1999;   99WO-US030270.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Frudakis TN;
XX
DR   WPI; 2000-431589/37.
XX
PT   Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT   encoding it, useful for the diagnosis, prevention and treatment of
PT   cancer, preferably ovarian cancer.
XX
PS   Claim 18; Fig 15; 299pp; English.
XX
CC   The present invention describes an isolated polypeptide comprising an
CC   immunogenic portion of an ovarian carcinoma protein (or its variants).
CC   Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC   cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC   carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC   for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC   cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557 represent human
CC   ovarian carcinoma polynucleotides and proteins used in the
CC   exemplification of the present invention
XX
SQ   Sequence 518 BP; 141 A; 132 C; 128 G; 116 T; 0 U; 1 Other;

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 52
ABN72736
ID   ABN72736 standard; DNA; 518 BP.
XX
AC   ABN72736;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Ovarian carcinoma antigen polynucleotide #152.
XX
KW   Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS   Homo sapiens.
XX
PN   WO200206317-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-US022635.
XX
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
PR   04-APR-2001; 2001US-00827271.
PR   18-JUN-2001; 2001US-00884441.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Fling SP,  Retter MW,  Panger GR;
PI   Reed SG,  Vedvick TS,  Carter D,  Hill P,  Albone E;
XX
DR   WPI; 2002-164781/21.
XX
PT   Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT   protein or its variants, useful for stimulating an immune response in a
PT   patient and treating ovarian cancer.
XX
PS   Example 1; Page 145; 408pp; English.
XX
CC   This invention relates to polypeptides comprising an immunogenic portion
CC   of an ovarian carcinoma protein which acts as an immunostimulant and is
CC   cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC   proteins, T cell populations and antigen presenting cells that express
CC   the polypeptides are useful for stimulating an immune response in a
CC   patient and treating ovarian cancer. This sequence represents DNA related
CC   to the invention
XX
SQ   Sequence 518 BP; 141 A; 132 C; 128 G; 116 T; 0 U; 1 Other;
```

```
  Query Match             16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 53
ADA08901
ID   ADA08901 standard; DNA; 518 BP.
XX
AC   ADA08901;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human ovarian carcinoma antigen polynucleotide #152.
XX
KW   ds; human; gene therapy; ovarian cancer; cancer.
XX
OS   Homo sapiens.
XX
PN   US2003091580-A1.
XX
PD   15-MAY-2003.
XX
PF   17-JUL-2001; 2001US-00907969.
XX
PR   18-JUN-2001; 2001US-00884441.
XX
PA   (MITC/) MITCHAM J L.
PA   (KING/) KING G E.
PA   (ALGA/) ALGATE P A.
PA   (FLIN/) FLING S P.
PA   (RETT/) RETTER M W.
PA   (FANG/) FANGER G R.
PA   (REED/) REED S G.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (HILL/) HILL P.
PA   (ALBO/) ALBONE E.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Fling SP,  Retter MW,  Fanger GR;
PI   Reed SG,  Vedvick TS,  Carter D,  Hill P,  Albone E;
XX
DR   WPI; 2003-532352/50.
XX
PT   New isolated O772P polypeptides and polynucleotides, useful in gene
PT   therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT   cancer.
XX
PS   Example 1; Fig 15; 371pp; English.
XX
CC   The invention relates to an isolated O772P polypeptide, which has the
CC   structure fully defined in the specification. The composition containing
CC   the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC   or antigen presenting cells are useful for stimulating an immune response
CC   and treating ovarian cancer. Detecting the presence of the
CC   polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC   carcinoma cDNAs and protein cDNAs were identified using microarray
CC   technology. The present sequence represents a human ovarian carcinoma
CC   antigen polynucleotide.
XX
SQ   Sequence 518 BP; 141 A; 132 C; 128 G; 116 T; 0 U; 1 Other;

  Query Match             16.9%;  Score 370;  DB 9;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 54
ADF08648
ID   ADF08648 standard; cDNA; 518 BP.
XX
AC   ADF08648;
XX
DT   12-FEB-2004  (first entry)
XX
DE   cDNA encoding secreted ovarian carcinoma antigen seqid 152.
XX
KW   gene therapy; protein therapy; vaccine; antibody inhibition;
KW   breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW   secreted ovarian carcinoma antigen; gene; ss.
XX
OS   Homo sapiens.
```

XX
PN US2003124140-A1.
XX
PD 03-JUL-2003.
XX
PF 17-JUL-2002; 2002US-00198053.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2003-897152/82.
XX
PT Oncogenic nucleic acids useful for the prevention, diagnosis and
PT treatment of breast cancer.
XX
PS Example 1; SEQ ID NO 152; 399pp; English.
XX
CC The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of (II) by expressing
CC inactive proteins or to supplement the patients own production of (II).
CC Additionally, (I) may be used to produce (II), by inserting (I) into a
CC host cell and culturing the cell to express the protein (II). (I) And its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The host cell may also be used as antigens in the production of
CC antibodies against (II) and in assays to identify modulators of (II)'s
CC expression and activity. The anti-(II) antibodies, agonists and
CC antagonists may be used to regulate expression and activity and as
CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC immunoassay). This sequence represents a polynucleotide encoding a
CC secreted ovarian carcinoma antigen.
XX
SQ Sequence 518 BP; 141 A; 132 C; 128 G; 116 T; 0 U; 1 Other;

Query Match 16.9%; Score 370; DB 10; Length 518;
Best Local Similarity 99.6%; Pred. No. 3.3e-173;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 55
ADG46396
ID ADG46396 standard; cDNA; 518 BP.
XX
AC ADG46396;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human ovarian carcinoma polynucleotide #152.
XX
KW Human; ovarian carcinoma; gene; ss; O8E; ovarian cancer;
KW secreted tumour antigen; cytostatic; O772P.
XX
OS Homo sapiens.
XX
PN US2003165504-A1.
XX
PD 04-SEP-2003.
XX
PF 04-APR-2001; 2001US-00827271.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
XX
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
XX
PI Retter MW, Fanger GR;
XX
DR WPI; 2003-898035/82.
XX
PT New isolated O8E or O772P polypeptides, useful for diagnosing,
PT preventing, treating and monitoring cancer, e.g. ovarian cancer,
PT stimulating the immune response in patient.
XX
PS Example 1; SEQ ID NO 152; 290pp; English.
XX
CC The invention relates to human ovarian carcinoma polypeptides, designated
CC O8E or O772P, and the polynucleotides encoding them. The invention also
CC relates to methods for inhibiting the development of cancer, e.g. ovarian
CC cancer in a patient, methods for stimulating and/or expanding T cells and
CC methods for identifying secreted tumour antigens. The polypeptides,
CC compositions, antibodies to the polypeptides and methods are useful for
CC diagnosing, preventing, treating and monitoring cancer, e.g. ovarian
CC cancer. The composition is particularly useful for stimulating an immune
CC response in patient. This sequence represents a human ovarian carcinoma
CC polynucleotide of the invention.
XX
SQ Sequence 518 BP; 141 A; 132 C; 128 G; 116 T; 0 U; 1 Other;

Query Match 16.9%; Score 370; DB 10; Length 518;
Best Local Similarity 99.6%; Pred. No. 3.3e-173;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 56
ADX17531
ID ADX17531 standard; cDNA; 518 BP.
XX
AC ADX17531;
XX
DT 21-APR-2005 (first entry)
XX
DE Human ovarian carcinoma antigen cDNA #145.
XX
KW diagnosis; cytostatic; immunostimulant; gene therapy; tumor;
KW ovarian tumor; cancer; carcinoma; antigen; ss.
XX
OS Homo sapiens.
XX
PN US2005031634-A1.
XX
PD 10-FEB-2005.
XX
PF 02-JUN-2004; 2004US-00860790.
XX
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
PR 17-JUL-2002; 2002US-00198053.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2005-151645/16.
XX
PT New ovarian carcinoma polynucleotides, preferably cDNAs, useful for
PT diagnosing, preventing and treating diseases, such as ovarian cancer, and
PT for eliciting humoral and/or cellular immune response.
XX
PS Example 1; SEQ ID NO 152; 398pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising a sequence
CC of, a sequence hybridizing under highly stringent conditions to, or
CC having at least 75, or 90% identity to a 849 or 1399 bp, fully defined in
CC the specification, its complement or degenerate variants, or a sequence
CC of at least 20 contiguous residues of the 849 or 1399 bp sequence. The
CC polynucleotides and polypeptides are useful for diagnosing, preventing
CC and treating diseases, such as ovarian cancer, and for eliciting humoral
CC and/or cellular immune response. This sequence corresponds to an ovarian
CC carcinoma antigen cDNA of the invention.
XX
SQ Sequence 518 BP; 141 A; 132 C; 128 G; 116 T; 0 U; 1 Other;

```
  Query Match             16.9%;  Score 370;  DB 14;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 57
ACF88710/c
ID ACF88710 standard; DNA; 419 BP.
XX
AC ACF88710;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human SIRS/sepsis diagnostic marker DNA fragment 7570.
XX
KW Systemic inflammatory response syndrome; SIRS; antibacterial;
KW immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS Homo sapiens.
XX
PN WO2004087949-A2.
XX
PD 14-OCT-2004.
XX

PF 31-MAR-2004; 2004WO-EP003419.
XX
PR 02-APR-2003; 2003DE-01015031.
PR 08-AUG-2003; 2003DE-01036511.
PR 02-SEP-2003; 2003DE-01040395.
XX
PA (SIRS-) SIRS LAB GMBH.
XX
PI Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;
XX
DR WPI; 2004-748070/73.
XX
PT In vitro detection of systemic inflammatory response syndrome and related
PT conditions, for e.g. monitoring progression, comprises detecting abnormal
PT expression of disease-related genes.
XX
PS Disclosure; Page; 75pp; German.
XX
CC The invention relates to a novel method for in vitro detection of
CC systemic inflammatory response syndrome (SIRS). The method comprises
CC detecting abnormal expression of disease-related genes, or their
CC associated peptides. The method of the invention demonstrates
CC antibacterial, immunosuppressive and antiinflammatory applications and
CC may be used for early differential diagnosis, monitoring progression,
CC assessing risk, assessing the likely response to treatment and for post
CC mortem diagnosis of systemic inflammatory response syndrome, sepsis and
CC sepsis-like conditions. The recombinant or synthetic nucleic acid
CC sequences of the invention, or derived proteins or peptides, may be
CC useful as calibrants in assays for the specified diseases, for evaluating
CC activity or toxicity in screening for active agents and/or for
CC preparation of agents for treatment or prevention of the specified
CC diseases. The current sequence is that of a human SIRS/sepsis diagnostic
CC marker DNA fragment of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at ftp.wipo.int/pub/published
CC pct_sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
CC disclosed within the specification, however, these have not been taken
CC into account during indexing due to inconsistencies in application and
CC format
XX
SQ Sequence 419 BP; 97 A; 100 C; 105 G; 116 T; 0 U; 1 Other;

Query Match 14.0%; Score 307; DB 13; Length 419;
Best Local Similarity 100.0%; Pred. No. 6.6e-142;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1886 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 1945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 259

Qy       1946 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 2005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 199

Qy       2006 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 2065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 139

Qy       2066 CTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTA 2125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 CTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTA 79

Qy       2126 AGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTG 2185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 AGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTG 19

Qy       2186 CATTCCC 2192
             |||||||
Db          18 CATTCCC 12
```

RESULT 58
ACD92815

ID ACD92815 standard; cDNA; 400 BP.
XX
AC ACD92815;
XX
DT 23-SEP-2003 (first entry)
XX
DE Human colon cancer cell expressed cDNA #1227.
XX
KW Open reading frame detection; genome sequencing; colon cancer;
KW breast cancer; population genome analysis; genetic shift; cancer;
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
KW agriculture; food crop genome; resistance gene; retrovirus;
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN US2002155438-A1.
XX
PD 24-OCT-2002.
XX
PF 27-SEP-1999; 99US-00406117.
XX
PR 20-NOV-1998; 98US-00196716.
XX
PA (SIMP/) SIMPSON A J G.
PA (NETO/) NETO E D.
PA (BREN/) BRENTANI R R.
XX
PI Simpson AJG, Neto ED, Brentani RR;
XX
DR WPI; 2003-182626/18.
XX
PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.
XX
PS Example 9; Page 190; 959pp; English.
XX
CC The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA , sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame; for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a foetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library

```
XX
SQ   Sequence 400 BP; 112 A; 100 C; 111 G; 76 T; 0 U; 1 Other;

  Query Match             13.8%;  Score 303;  DB 10;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-140;
  Matches  303;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1517 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 145

Qy       1577 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1636
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 205

Qy       1637 GCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1696
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 265

Qy       1697 GACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA 1756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA 325

Qy       1757 AGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 1816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 385

Qy       1817 AGA 1819
             |||
Db        386 AGA 388


RESULT 59
AAI29110
ID   AAI29110 standard; cDNA; 412 BP.
XX
AC   AAI29110;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Colon tumour related determined cDNA sequence for clone CT2-112.
XX
KW   Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
KW   gene therapy; vaccine; colonic cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200149716-A2.
XX
PD   12-JUL-2001.
XX
PF   29-DEC-2000; 2000WO-US035596.
XX
PR   30-DEC-1999;   99US-00476296.
PR   10-JAN-2000; 2000US-00480321.
PR   15-FEB-2000; 2000US-00504629.
PR   06-MAR-2000; 2000US-00519444.
PR   19-MAY-2000; 2000US-00575251.
PR   29-JUN-2000; 2000US-00609448.
PR   28-AUG-2000; 2000US-00649811.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Lodes MJ,  Secrist H,  Benson DR,  Meagher MJ,  Stolk JA;
PI   King GE,  Wang T,  Jiang Y;
XX
DR   WPI; 2001-441847/47.
XX
PT   Colon tumor associated proteins and nucleic acids useful for the
PT   prevention, diagnosis and treatment of colonic cancer.
XX
PS   Claim 25; Page 311; 472pp; English.
XX
CC   The present invention describes colon tumour associated proteins (I) and
CC   the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC   (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC   (II) may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate colon tumour associated protein (TCAP)
CC   expression, such as colonic cancer. For example, (I) and (II) may be used
CC   to treat disorders associated with decreased expression by rectifying
CC   mutations or deletions in a patient's genome that affect the activity of
CC   TCAPs by expressing inactive proteins or to supplement the patients own
CC   production of them. Additionally, (II) may be used to produce the TCAP
CC   proteins, by inserting the nucleic acids into a host cell culturing the
CC   cell to express the protein. (II) and its complementary sequences may
CC   also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
CC   and hybridisation assays to detect and quantitate the presence of similar
CC   nucleic acids in samples, and therefore which patients may be in need of
CC   restorative therapy. (I) may also be used as antigens in the production
CC   of antibodies against TCAPs and in assays to identify modulators of TCAP
CC   expression and activity. Anti-(I) antibodies and antagonists may also be
CC   used to down regulate TCAP expression and activity. The anti-(I)
CC   antibodies may also be used as diagnostic agents for detecting the
CC   presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay
CC   (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent
CC   nucleotide and amino acid sequences given in the exemplification of the
CC   present invention
XX
SQ   Sequence 412 BP; 113 A; 109 C; 109 G; 81 T; 0 U; 0 Other;

  Query Match             13.8%;  Score 303;  DB 4;  Length 412;
  Best Local Similarity   99.7%;  Pred. No. 6.4e-140;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 60
ABZ33296
ID   ABZ33296 standard; cDNA; 412 BP.
XX
AC   ABZ33296;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Human colon tumour cDNA for clone CT2-112 SEQ ID NO:660.
XX
KW   Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
KW   tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200283070-A2.
XX
```

```
PD   24-OCT-2002.
XX
PF   09-APR-2002; 2002WO-US011475.
XX
PR   10-APR-2001; 2001US-00833263.
PR   03-AUG-2001; 2001US-00922217.
PR   19-DEC-2001; 2001US-00025380.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Lodes MJ,  Secrist H,  Benson DR,  Meagher MJ,  Stolk JA;
PI   Wang T,  Jiang Y,  Smith CL,  King GE,  Wang A,  Clapper JD,  Skeiky YAW;
PI   Fanger GR,  Vedvick TS,  Carter D;
XX
DR   WPI; 2003-067548/06.
XX
PT   New polynucleotide, useful for the preparation of a composition for
PT   stimulating an immune response against, or treating, cancer.
XX
PS   Example 3; Page 314; 537pp; English.
XX
CC   The present invention describes compounds (I) for the immunotherapy and
CC   diagnosis of colon cancer. Also described: (1) a method for detecting the
CC   presence of cancer in a patient; (2) a method for stimulating and/or
CC   expanding T cells specific for a tumour protein; (3) an isolated T cell
CC   population comprising T cells prepared by the method of (2); (4) a method
CC   for stimulating an immune response in a patient; (5) a method for
CC   treating cancer in a patient; and (6) a method for inhibiting the
CC   development of cancer in a patient. (I) have immunostimulant and
CC   cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
CC   and ABP55343 to ABP55391 represent human colon cancer/tumour related
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 412 BP; 113 A; 109 C; 109 G; 81 T; 0 U; 0 Other;

  Query Match           13.8%;  Score 303;  DB 8;  Length 412;
  Best Local Similarity   99.7%;  Pred. No. 6.4e-140;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 61
ACH22349
ID   ACH22349 standard; cDNA; 482 BP.
XX
AC   ACH22349;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult ovary cDNA #729.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 9561; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 482 BP; 128 A; 121 C; 126 G; 102 T; 0 U; 5 Other;

  Query Match           13.7%;  Score 300;  DB 9;  Length 482;
  Best Local Similarity   99.7%;  Pred. No. 2e-138;
  Matches  350;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        106 TCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 TCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 191

Qy        166 AGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATG 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 AGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATG 251

Qy        226 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 311

Qy        286 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 371

Qy        346 TGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTC 405
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        372 TCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTC 431

Qy        406 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 482


RESULT 62
ACH16408
ID   ACH16408 standard; cDNA; 393 BP.
XX
AC   ACH16408;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult heart cDNA #722.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 3620; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 393 BP; 115 A; 94 C; 92 G; 92 T; 0 U; 0 Other;

  Query Match           13.4%;  Score 294;  DB 9;  Length 393;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-135;
  Matches  294;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         45 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 104

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 164

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 224

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 284

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 338


RESULT 63
ACH20515
ID   ACH20515 standard; cDNA; 399 BP.
XX
AC   ACH20515;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult liver cDNA #127.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 7727; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
```

```
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 399 BP; 106 A; 114 C; 102 G; 77 T; 0 U; 0 Other;

  Query Match            12.5%;  Score 275;  DB 9;  Length 399;
  Best Local Similarity  100.0%;  Pred. No. 5.2e-126;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 106

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 166

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 226

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 286

Qy        456 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
              |||||||||||||||||||||||||||||||||||
Db        287 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 321


RESULT 64
AAX23500
ID   AAX23500 standard; DNA; 284 BP.
XX
AC   AAX23500;
XX
DT   17-JUN-1999  (first entry)
XX
DE   Human neutrophil cDNA clone 927.
XX
KW   Neutrophil; gene expression profile; granulocyte; pathogen-exposed;
KW   sterile inflammatory disease; detection; therapeutic agent; human;
KW   expression modulator; pathogenic infection; cell activation; primer;
KW   global transcriptional response; ss.
XX
OS   Homo sapiens.
XX
PN   WO9910536-A1.
XX
PD   04-MAR-1999.
XX
PF   21-AUG-1998;   98WO-US017284.
XX
PR   22-AUG-1997;   97US-0056844P.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   Yerramilli SV,  Prashar Y,  Newburger P,  Goguen J,  Weissman SM;
XX
DR   WPI; 1999-204678/17.
XX
PT   Detection of pathogen exposure or sterile inflammatory disease in a
PT   subject - by comparing gene expression profiles from granulocytes of the
PT   patient and control granulocytes.
XX
PS   Example 3; Page 54; 84pp; English.
XX
CC   This invention describes a method for the comparison of gene expression
CC   profiles from granulocytes from a test subject and from pathogen-exposed
CC   or sterile inflammatory disease granulocytes or quiescent granulocytes.
CC   The method is used to detect pathogen exposure or sterile inflammatory
CC   disease in a subject and to identify therapeutic agents that modulate
CC   expression of a gene in response to a pathogenic infection or to sterile
CC   inflammatory disease in a subject. The method tests for global
CC   transcriptional response of granulocytes during cell activation
XX
SQ   Sequence 284 BP; 76 A; 69 C; 69 G; 70 T; 0 U; 0 Other;

  Query Match            12.3%;  Score 270;  DB 2;  Length 284;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-123;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 74

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 134

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 194

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 254

Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTG 2182
              |||||||||||||||||||||||||||||
Db        255 TTTCATTGCACTTTCCTTTCAATAAAGTTG 284


RESULT 65
ACH20853
ID   ACH20853 standard; cDNA; 411 BP.
XX
AC   ACH20853;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult liver cDNA #465.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 8065; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
```

```
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 411 BP; 107 A; 104 C; 111 G; 89 T; 0 U; 0 Other;

  Query Match             11.9%;  Score 260;  DB 9;  Length 411;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-118;
  Matches  260;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 60

Qy         87 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 120

Qy        147 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 180

Qy        207 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 240

Qy        267 CTGCACCTGGTGCTCCGTCT 286
              ||||||||||||||||||||
Db        241 CTGCACCTGGTGCTCCGTCT 260


RESULT 66
ADU14740
ID   ADU14740 standard; DNA; 543 BP.
XX
AC   ADU14740;
XX
DT   27-JAN-2005  (first entry)
XX
DE   Solid tumour prognosis gene seqid 5179.
XX
KW   cytostatic; gene therapy; expression profile; solid tumour;
KW   peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
OS   Unidentified.
XX
PN   WO2004097052-A2.
XX
PD   11-NOV-2004.
XX
PF   29-APR-2004; 2004WO-US013587.
XX
PR   29-APR-2003; 2003US-0466067P.
PR   23-JAN-2004; 2004US-0538246P.
XX
PA   (AMHP ) WYETH.
PA   (STRA/) STRAHS A.
XX
PI   Strahs A,  Trepicchio WL,  Burczynski ME,  Twine NC,  Slonim DK;
PI   Immerman F,  Dorner AJ;
XX
DR   WPI; 2004-804779/79.
XX
PT   A method, useful for prognosing and treating solid tumor, comprises
PT   comparing an expression profile of a gene expressed in peripheral blood
PT   mononuclear cells to a reference expression profile of a gene.
XX
PS   Disclosure; Page; 111pp; English.
XX
CC   The invention describes a method comprising comparing an expression
CC   profile of at least one gene in a peripheral blood sample of a patient to
CC   at least one reference expression profile of the at least one gene, where
CC   the patient has a solid tumour, and each of the gene is differentially
CC   expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC   of patients as compared to PBMCs of a second class of patients, where
CC   both the first and second classes of patients have the solid tumour, and
CC   each of the first and second classes is a subcluster formed by an
CC   unsupervised clustering analysis of gene expression profiles in PBMCs of
CC   a population of patients who have the solid tumour, and where the
CC   majority of the first class of patients has a first clinical outcome, and
CC   the majority of the second class of patients has a second clinical
CC   outcome. Also described are: a system comprising (i) a memory or a
CC   storage medium including data that represent an expression profile of at
CC   least one gene in a peripheral blood sample of a patient who has a solid
CC   tumour, (ii) at least another storage medium including data that
CC   represent at least one reference expression profile of the gene, (iii) a
CC   program capable of comparing the expression profile to the reference
CC   expression profile, and (iv) a processor capable of executing the
CC   program, where expression levels of the gene in peripheral blood
CC   mononuclear cells of patients who have the solid tumour correlate with
CC   clinical outcomes of the patients; and a nucleic acid or protein array
CC   comprising concentrated probes for solid tumour prognosis genes, where
CC   each of the solid tumour prognosis genes is differentially expressed in
CC   PBMCs of a first class of patients as compared to PBMCs of a second class
CC   of patients, where both the first and second classes of patients have a
CC   solid tumour, and where the first class of patients has a first clinical
CC   outcome, and the second class of patients has a second clinical outcome.
CC   The method, system, and array are useful for prognosing and treating
CC   solid tumours. This sequence represents a solid tumour prognosis gene of
CC   the invention. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 543 BP; 128 A; 128 C; 126 G; 107 T; 0 U; 54 Other;

  Query Match             11.6%;  Score 255;  DB 13;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-116;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1919 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 1978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 348

Qy       1979 GATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 2038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 GATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 408

Qy       2039 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTT 2098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTT 468

Qy       2099 GGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCAT 2158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCAT 528

Qy       2159 TGCACTTTCCTTTCA 2173
              ||||||||||||||||
Db        529 TGCACTTTCCTTTCA 543


RESULT 67
ADU14804
ID   ADU14804 standard; DNA; 565 BP.
XX
```

AC ADU14804;
XX
DT 27-JAN-2005 (first entry)
XX
DE Solid tumour prognosis gene seqid 5243.
XX
KW cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
OS Unidentified.
XX
PN WO2004097052-A2.
XX
PD 11-NOV-2004.
XX
PF 29-APR-2004; 2004WO-US013587.
XX
PR 29-APR-2003; 2003US-0466067P.
PR 23-JAN-2004; 2004US-0538246P.
XX
PA (AMHP) WYETH.
PA (STRA/) STRAHS A.
XX
PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dorner AJ;
XX
DR WPI; 2004-804779/79.
XX
PT A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
PS Disclosure; Page; 111pp; English.
XX
CC The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 565 BP; 150 A; 150 C; 146 G; 119 T; 0 U; 0 Other;

Query Match 11.6%; Score 254; DB 13; Length 565;
Best Local Similarity 99.7%; Pred. No. 1.4e-115;
Matches 304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1860 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 320

Qy       1920 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 1979
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GGTAAGACCATTACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG.380

Qy       1980 ATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 440

Qy       2040 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTG 2099
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTG 500

Qy       2100 GTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATT 2159
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATT 560

Qy       2160 GCACT 2164
             |||||
Db        561 GCACT 565
```

RESULT 68
AAA69735
ID AAA69735 standard; cDNA; 585 BP.
XX
AC AAA69735;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide SEQ ID NO:45.
XX
KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200036107-A2.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US030270.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
DR WPI; 2000-431589/37.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer.
XX
PS Claim 1; Fig 1; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557 represent human
CC ovarian carcinoma polynucleotides and proteins used in the
CC exemplification of the present invention

```
XX
SQ   Sequence 585 BP; 153 A; 152 C; 145 G; 126 T; 0 U; 9 Other;

  Query Match            11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-112;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 69
ABN72629
ID   ABN72629 standard; DNA; 585 BP.
XX
AC   ABN72629;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Ovarian carcinoma antigen polynucleotide #45.
XX
KW   Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS   Homo sapiens.
XX
PN   WO200206317-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-US022635.
XX
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
PR   04-APR-2001; 2001US-00827271.
PR   18-JUN-2001; 2001US-00884441.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Fling SP,  Retter MW,  Fanger GR;
PI   Reed SG,  Vedvick TS,  Carter D,  Hill P,  Albone E;
XX
DR   WPI; 2002-164781/21.
XX
PT   Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT   protein or its variants, useful for stimulating an immune response in a
PT   patient and treating ovarian cancer.
XX
PS   Example 1; Page 105; 408pp; English.
XX
CC   This invention relates to polypeptides comprising an immunogenic portion
CC   of an ovarian carcinoma protein which acts as an immunostimulant and is
CC   cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC   proteins, T cell populations and antigen presenting cells that express
CC   the polypeptides are useful for stimulating an immune response in a
CC   patient and treating ovarian cancer. This sequence represents DNA related
CC   to the invention
XX
SQ   Sequence 585 BP; 153 A; 152 C; 145 G; 126 T; 0 U; 9 Other;

  Query Match            11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-112;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 70
ADA08794
ID   ADA08794 standard; DNA; 585 BP.
XX
AC   ADA08794;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human ovarian carcinoma antigen polynucleotide #45.
XX
KW   ds; human; gene therapy; ovarian cancer; cancer.
XX
OS   Homo sapiens.
XX
PN   US2003091580-A1.
XX
PD   15-MAY-2003.
XX
PF   17-JUL-2001; 2001US-00907969.
XX
PR   18-JUN-2001; 2001US-00884441.
XX
PA   (MITC/) MITCHAM J L.
PA   (KING/) KING G E.
PA   (ALGA/) ALGATE P A.
PA   (FLIN/) FLING S P.
PA   (RETT/) RETTER M W.
PA   (FANG/) FANGER G R.
PA   (REED/) REED S G.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (HILL/) HILL P.
PA   (ALBO/) ALBONE E.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Fling SP,  Retter MW,  Fanger GR;
PI   Reed SG,  Vedvick TS,  Carter D,  Hill P,  Albone E;
XX
DR   WPI; 2003-532352/50.
XX
PT   New isolated O772P polypeptides and polynucleotides, useful in gene
PT   therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT   cancer.
XX
PS   Example 1; Fig 1; 371pp; English.
```

```
XX
CC   The invention relates to an isolated O772P polypeptide, which has the
CC   structure fully defined in the specification. The composition containing
CC   the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC   or antigen presenting cells are useful for stimulating an immune response
CC   and treating ovarian cancer. Detecting the presence of the
CC   polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC   carcinoma cDNAs and protein cDNAs were identified using microarray
CC   technology. The present sequence represents a human ovarian carcinoma
CC   antigen polynucleotide.
XX
SQ   Sequence 585 BP; 153 A; 152 C; 145 G; 126 T; 0 U; 9 Other;

  Query Match             11.3%;  Score 248;  DB 9;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-112;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 71
ADF08541
ID   ADF08541 standard; cDNA; 585 BP.
XX
AC   ADF08541;
XX
DT   12-FEB-2004  (first entry)
XX
DE   cDNA encoding secreted ovarian carcinoma antigen seqid 45.
XX
KW   gene therapy; protein therapy; vaccine; antibody inhibition;
KW   breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW   secreted ovarian carcinoma antigen; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003124140-A1.
XX
PD   03-JUL-2003.
XX
PF   17-JUL-2002; 2002US-00198053.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
PR   04-APR-2001; 2001US-00827271.
PR   18-JUN-2001; 2001US-00884441.
PR   17-JUL-2001; 2001US-00907969.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Bangur CS,  Retter MW,  Fanger GR,  Hill P;
XX
DR   WPI; 2003-897152/82.
XX
PT   Oncogenic nucleic acids useful for the prevention, diagnosis and
PT   treatment of breast cancer.
XX
PS   Example 1; SEQ ID NO 45; 399pp; English.
XX
CC   The invention describes nucleic acids (I) and the polypeptides (II) they
CC   encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC   treating diseases related to their aberrant expression i.e. breast
CC   cancers. For example, (I) and (II) may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of (II) by expressing
CC   inactive proteins or to supplement the patients own production of (II).
CC   Additionally, (I) may be used to produce (II), by inserting (I) into a
CC   host cell and culturing the cell to express the protein (II). (I) And its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids in
CC   samples, and therefore which patients may be in need of restorative
CC   therapy. The host cell may also be used as antigens in the production of
CC   antibodies against (II) and in assays to identify modulators of (II)'s
CC   expression and activity. The anti-(II) antibodies, agonists and
CC   antagonists may be used to regulate expression and activity and as
CC   diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC   immunoassay). This sequence represents a polynucleotide encoding a
CC   secreted ovarian carcinoma antigen.
XX
SQ   Sequence 585 BP; 153 A; 152 C; 145 G; 126 T; 0 U; 9 Other;

  Query Match             11.3%;  Score 248;  DB 10;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-112;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 72
ADG46289
ID   ADG46289 standard; cDNA; 585 BP.
XX
AC   ADG46289;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human ovarian carcinoma polynucleotide #45.
XX
KW   Human; ovarian carcinoma; gene; ss; O8E; ovarian cancer;
KW   secreted tumour antigen; cytostatic; O772P.
XX
OS   Homo sapiens.
XX
PN   US2003165504-A1.
XX
```

PD 04-SEP-2003.
XX
PF 04-APR-2001; 2001US-00827271.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
XX
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
XX
PI Retter MW, Fanger GR;
XX
DR WPI; 2003-898035/82.
XX
PT New isolated O8E or O772P polypeptides, useful for diagnosing,
PT preventing, treating and monitoring cancer, e.g. ovarian cancer,
PT stimulating the immune response in patient.
XX
PS Example 1; SEQ ID NO 45; 290pp; English.
XX
CC The invention relates to human ovarian carcinoma polypeptides, designated
CC O8E or O772P, and the polynucleotides encoding them. The invention also
CC relates to methods for inhibiting the development of cancer, e.g. ovarian
CC cancer in a patient, methods for stimulating and/or expanding T cells and
CC methods for identifying secreted tumour antigens. The polypeptides,
CC compositions, antibodies to the polypeptides and methods are useful for
CC diagnosing, preventing, treating and monitoring cancer, e.g. ovarian
CC cancer. The composition is particularly useful for stimulating an immune
CC response in patient. This sequence represents a human ovarian carcinoma
CC polynucleotide of the invention.
XX
SQ Sequence 585 BP; 153 A; 152 C; 145 G; 126 T; 0 U; 9 Other;

Query Match 11.3%; Score 248; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
Db 292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy 1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
Db 352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy 2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
Db 412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy 2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
Db 472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy 2139 AAGGTTTC 2146
Db 532 AAGGTTTC 539

RESULT 73
ADX17424
ID ADX17424 standard; cDNA; 585 BP.
XX
AC ADX17424;
XX
DT 21-APR-2005 (first entry)
XX
DE Human ovarian carcinoma antigen cDNA #45.
XX
KW diagnosis; cytostatic; immunostimulant; gene therapy; tumor;
KW ovarian tumor; cancer; carcinoma; antigen; ss.
XX
OS Homo sapiens.
XX
PN US2005031634-A1.
XX
PD 10-FEB-2005.
XX
PF 02-JUN-2004; 2004US-00860790.
XX
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
PR 17-JUL-2002; 2002US-00198053.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2005-151645/16.
XX
PT New ovarian carcinoma polynucleotides, preferably cDNAs, useful for
PT diagnosing, preventing and treating diseases, such as ovarian cancer, and
PT for eliciting humoral and/or cellular immune response.
XX
PS Example 1; SEQ ID NO 45; 398pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising a sequence
CC of, a sequence hybridizing under highly stringent conditions to, or
CC having at least 75, or 90% identity to a 849 or 1399 bp, fully defined in
CC the specification, its complement or degenerate variants, or a sequence
CC of at least 20 contiguous residues of the 849 or 1399 bp sequence. The
CC polynucleotides and polypeptides are useful for diagnosing, preventing
CC and treating diseases, such as ovarian cancer, and for eliciting humoral
CC and/or cellular immune response. This sequence corresponds to an ovarian
CC carcinoma antigen cDNA of the invention.
XX
SQ Sequence 585 BP; 153 A; 152 C; 145 G; 126 T; 0 U; 9 Other;

Query Match 11.3%; Score 248; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
Db 292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy 1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
Db 352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy 2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
Db 412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy 2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
Db 472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy 2139 AAGGTTTC 2146
Db 532 AAGGTTTC 539

RESULT 74
ABV45876/c

```
ID   ABV45876 standard; cDNA; 467 BP.
XX
AC   ABV45876;
XX
DT   16-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 45867.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 9067; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 467 BP; 94 A; 133 C; 121 G; 119 T; 0 U; 0 Other;

  Query Match             11.1%;  Score 244;  DB 5;  Length 467;
  Best Local Similarity   99.4%;  Pred. No. 1.3e-110;
  Matches  344;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        343 AGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCA 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCA 408

Qy        403 TTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 462
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TTCCTCTTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 348

Qy        463 TGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 522
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        347 TGTCTGACTACAACATCCAGAAAGAGTTTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 288

Qy        523 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 228

Qy        583 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATC 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATC 168

Qy        643 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        167 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 122


RESULT 75
ACF88709
ID   ACF88709 standard; DNA; 558 BP.
XX
AC   ACF88709;
XX
DT   02-JUN-2005  (first entry)
XX
DE   Human SIRS/sepsis diagnostic marker DNA fragment 7569.
XX
KW   Systemic inflammatory response syndrome; SIRS; antibacterial;
KW   immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX
OS   Homo sapiens.
XX
PN   WO2004087949-A2.
XX
PD   14-OCT-2004.
XX
PF   31-MAR-2004; 2004WO-EP003419.
XX
PR   02-APR-2003; 2003DE-01015031.
PR   08-AUG-2003; 2003DE-01036511.
PR   02-SEP-2003; 2003DE-01040395.
XX
PA   (SIRS-) SIRS LAB GMBH.
XX
PI   Russwurm S,  Reinhart K,  Saluz H,  Straube E,  Zipfel PF,  Deigner H;
XX
DR   WPI; 2004-748070/73.
XX
PT   In vitro detection of systemic inflammatory response syndrome and related
PT   conditions, for e.g. monitoring progression, comprises detecting abnormal
PT   expression of disease-related genes.
XX
PS   Disclosure; Page; 75pp; German.
XX
CC   The invention relates to a novel method for in vitro detection of
CC   systemic inflammatory response syndrome (SIRS). The method comprises
CC   detecting abnormal expression of disease-related genes, or their
CC   associated peptides. The method of the invention demonstrates
CC   antibacterial, immunosuppressive and antiinflammatory applications and
CC   may be used for early differential diagnosis, monitoring progression,
CC   assessing risk, assessing the likely response to treatment and for post
CC   mortem diagnosis of systemic inflammatory response syndrome, sepsis and
CC   sepsis-like conditions. The recombinant or synthetic nucleic acid
CC   sequences of the invention, or derived proteins or peptides, may be
CC   useful as calibrants in assays for the specified diseases, for evaluating
CC   activity or toxicity in screening for active agents and/or for
CC   preparation of agents for treatment or prevention of the specified
CC   diseases. The current sequence is that of a human SIRS/sepsis diagnostic
CC   marker DNA fragment of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at ftp.wipo.int/pub/published
CC   pct_sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
CC   disclosed within the specification, however, these have not been taken
CC   into account during indexing due to inconsistencies in application and
CC   format
XX
SQ   Sequence 558 BP; 143 A; 131 C; 155 G; 128 T; 0 U; 1 Other;

  Query Match             11.1%;  Score 243;  DB 13;  Length 558;
  Best Local Similarity   99.7%;  Pred. No. 4.1e-110;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;
```

```
Qy        835 CAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 60

Qy        895 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 120

Qy        955 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1014
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTGGT-CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 179

Qy       1015 CCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 CCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 239

Qy       1075 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 299

Qy       1135 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCC 1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCC 359

Qy       1195 GTCT 1198
              ||||
Db        360 GTCT 363
```

RESULT 76
AAH71976
ID AAH71976 standard; cDNA; 286 BP.
XX
AC AAH71976;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 3250.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 637-638; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 286 BP; 80 A; 74 C; 77 G; 55 T; 0 U; 0 Other;

```
  Query Match             11.0%;  Score 242;  DB 4;  Length 286;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-109;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 81

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 141

Qy        870 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 201

Qy        930 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 261

Qy        990 TT 991
              ||
Db        262 TT 263
```

RESULT 77
AAH70916
ID AAH70916 standard; cDNA; 487 BP.
XX
AC AAH70916;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 2190.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 463; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the

```
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 487 BP; 128 A; 123 C; 102 G; 129 T; 0 U; 5 Other;

  Query Match            11.0%;  Score 242;  DB 4;  Length 487;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-109;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 282

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 342

Qy        870 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 402

Qy        930 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 462

Qy        990 TT 991
              ||
Db        463 TT 464


RESULT 78
ABZ20057
ID   ABZ20057 standard; cDNA; 228 BP.
XX
AC   ABZ20057;
XX
DT   23-JAN-2003  (first entry)
XX
DE   Group III cDNA cancer related clone SEQ ID NO:2483.
XX
KW   Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW   immune response; virology; immunology; microbiology; molecular biology;
KW   recombinant DNA technology; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200278516-A2.
XX
PD   10-OCT-2002.
XX
PF   28-MAR-2002; 2002WO-US010421.
XX
PR   30-MAR-2001; 2001US-0280255P.
PR   28-AUG-2001; 2001US-0315563P.
PR   09-JAN-2002; 2002US-0347313P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang S,  Bangur CS,  Gaiger A;
XX
DR   WPI; 2003-058387/05.
XX
PT   New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT   preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT   virology, immunology, microbiology, molecular biology and recombinant DNA
PT   techniques.
XX
PS   Claim 1; SEQ ID NO 2483; 207pp; English.
XX
CC   ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC   ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC   invention. (I) and (II) have cytostatic activity and can be used in gene
CC   therapy and vaccines. (I), (II), antibodies and compositions from the
CC   present invention are useful for diagnosing, preventing and treating
CC   cancer, which expresses CT or CP mRNA antigens. They are useful for
CC   stimulating immune response. They can also be useful in virology,
CC   immunology, microbiology, molecular biology and recombinant DNA
CC   techniques. N.B. The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 228 BP; 63 A; 60 C; 61 G; 44 T; 0 U; 0 Other;

  Query Match            10.4%;  Score 228;  DB 8;  Length 228;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 60

Qy        504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 120

Qy        564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 180

Qy        624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 228


RESULT 79
ABZ19866
ID   ABZ19866 standard; cDNA; 228 BP.
XX
AC   ABZ19866;
XX
DT   23-JAN-2003  (first entry)
XX
DE   Group III cDNA cancer related clone SEQ ID NO:2292.
XX
KW   Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW   immune response; virology; immunology; microbiology; molecular biology;
KW   recombinant DNA technology; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200278516-A2.
XX
PD   10-OCT-2002.
XX
PF   28-MAR-2002; 2002WO-US010421.
XX
PR   30-MAR-2001; 2001US-0280255P.
PR   28-AUG-2001; 2001US-0315563P.
PR   09-JAN-2002; 2002US-0347313P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang S,  Bangur CS,  Gaiger A;
XX
DR   WPI; 2003-058387/05.
XX
PT   New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT   preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT   virology, immunology, microbiology, molecular biology and recombinant DNA
PT   techniques.
XX
PS   Claim 1; SEQ ID NO 2292; 207pp; English.
XX
CC   ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
```

CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 228 BP; 64 A; 59 C; 61 G; 44 T; 0 U; 0 Other;

Query Match 10.4%; Score 228; DB 8; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1356 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 60

Qy       1416 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 120

Qy       1476 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 180

Qy       1536 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 228
```

RESULT 80
ABZ19086/c
ID ABZ19086 standard; cDNA; 228 BP.
XX
AC ABZ19086;
XX
DT 23-JAN-2003 (first entry)
XX
DE Group III cDNA cancer related clone SEQ ID NO:1512.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US010421.
XX
PR 30-MAR-2001; 2001US-0280255P.
PR 28-AUG-2001; 2001US-0315563P.
PR 09-JAN-2002; 2002US-0347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX
PS Claim 1; SEQ ID NO 1512; 207pp; English.
XX

CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 228 BP; 44 A; 61 C; 59 G; 64 T; 0 U; 0 Other;

Query Match 10.4%; Score 228; DB 8; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1356 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 169

Qy       1416 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 109

Qy       1476 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 49

Qy       1536 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1
```

RESULT 81
ABZ18791
ID ABZ18791 standard; cDNA; 228 BP.
XX
AC ABZ18791;
XX
DT 23-JAN-2003 (first entry)
XX
DE Group III cDNA cancer related clone SEQ ID NO:1217.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US010421.
XX
PR 30-MAR-2001; 2001US-0280255P.
PR 28-AUG-2001; 2001US-0315563P.
PR 09-JAN-2002; 2002US-0347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX
PS Claim 1; SEQ ID NO 1217; 207pp; English.

```
XX
CC   ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC   ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC   invention. (I) and (II) have cytostatic activity and can be used in gene
CC   therapy and vaccines. (I), (II), antibodies and compositions from the
CC   present invention are useful for diagnosing, preventing and treating
CC   cancer, which expresses CT or CP mRNA antigens. They are useful for
CC   stimulating immune response. They can also be useful in virology,
CC   immunology, microbiology, molecular biology and recombinant DNA
CC   techniques. N.B. The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 228 BP; 64 A; 59 C; 61 G; 44 T; 0 U; 0 Other;

  Query Match             10.4%;  Score 228;  DB 8;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1356 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 60

Qy       1416 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 120

Qy       1476 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 180

Qy       1536 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 228


RESULT 82
ABZ19457
ID   ABZ19457 standard; cDNA; 228 BP.
XX
AC   ABZ19457;
XX
DT   23-JAN-2003  (first entry)
XX
DE   Group III cDNA cancer related clone SEQ ID NO:1883.
XX
KW   Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW   immune response; virology; immunology; microbiology; molecular biology;
KW   recombinant DNA technology; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200278516-A2.
XX
PD   10-OCT-2002.
XX
PF   28-MAR-2002; 2002WO-US010421.
XX
PR   30-MAR-2001; 2001US-0280255P.
PR   28-AUG-2001; 2001US-0315563P.
PR   09-JAN-2002; 2002US-0347313P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang S,  Bangur CS,  Gaiger A;
XX
DR   WPI; 2003-058387/05.
XX
PT   New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT   preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT   virology, immunology, microbiology, molecular biology and recombinant DNA
PT   techniques.
XX
PS   Claim 1; SEQ ID NO 1883; 207pp; English.
XX
CC   ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC   ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC   invention. (I) and (II) have cytostatic activity and can be used in gene
CC   therapy and vaccines. (I), (II), antibodies and compositions from the
CC   present invention are useful for diagnosing, preventing and treating
CC   cancer, which expresses CT or CP mRNA antigens. They are useful for
CC   stimulating immune response. They can also be useful in virology,
CC   immunology, microbiology, molecular biology and recombinant DNA
CC   techniques. N.B. The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 228 BP; 63 A; 61 C; 61 G; 43 T; 0 U; 0 Other;

  Query Match             10.4%;  Score 228;  DB 8;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1812 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 60

Qy       1872 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 120

Qy       1932 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 180

Qy       1992 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 228


RESULT 83
ABZ19715
ID   ABZ19715 standard; cDNA; 228 BP.
XX
AC   ABZ19715;
XX
DT   23-JAN-2003  (first entry)
XX
DE   Group III cDNA cancer related clone SEQ ID NO:2141.
XX
KW   Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW   immune response; virology; immunology; microbiology; molecular biology;
KW   recombinant DNA technology; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200278516-A2.
XX
PD   10-OCT-2002.
XX
PF   28-MAR-2002; 2002WO-US010421.
XX
PR   30-MAR-2001; 2001US-0280255P.
PR   28-AUG-2001; 2001US-0315563P.
PR   09-JAN-2002; 2002US-0347313P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang S,  Bangur CS,  Gaiger A;
XX
DR   WPI; 2003-058387/05.
XX
PT   New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT   preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT   virology, immunology, microbiology, molecular biology and recombinant DNA
PT   techniques.
```

```
XX
PS   Claim 1; SEQ ID NO 2141; 207pp; English.
XX
CC   ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC   ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC   invention. (I) and (II) have cytostatic activity and can be used in gene
CC   therapy and vaccines. (I), (II), antibodies and compositions from the
CC   present invention are useful for diagnosing, preventing and treating
CC   cancer, which expresses CT or CP mRNA antigens. They are useful for
CC   stimulating immune response. They can also be useful in virology,
CC   immunology, microbiology, molecular biology and recombinant DNA
CC   techniques. N.B. The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 228 BP; 64 A; 59 C; 61 G; 44 T; 0 U; 0 Other;

  Query Match             10.4%;  Score 228;  DB 8;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1356 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 60

Qy       1416 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 120

Qy       1476 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 180

Qy       1536 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 228


RESULT 84
ADR12560
ID   ADR12560 standard; DNA; 228 BP.
XX
AC   ADR12560;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Gene vaccine nucleic acid #14.
XX
KW   ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW   antibacterial; haemostatic; protozoacide; antiinflammatory;
KW   neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW   intracellular parasite; protozoan parasite infection; malaria;
KW   tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW   cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW   Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW   Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW   Herpes virus; severe acute respiratory syndrome.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .228
FT                   /*tag=  a
XX
PN   WO2004067040-A1.
XX
PD   12-AUG-2004.
XX
PF   30-JAN-2004; 2004WO-JP000975.
XX
PR   31-JAN-2003; 2003JP-00023507.
XX
PA   (KYUS-) KYUSHU TLO CO LTD.
XX
PI   Himeno K,  Ishii K;
XX
DR   WPI; 2004-594036/57.
DR   P-PSDB; ADR12561.
XX
PT   Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT   of intracellular parasite, useful for treating parasite infections such
PT   as malaria, tuberculosis, toxoplasmosis.
XX
PS   Disclosure; SEQ ID NO 27; 266pp; Japanese.
XX
CC   A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC   and a nucleic acid sequence encoding the antigen protein of an
CC   intracellular parasite containing T-cell target sequence. (I) is useful
CC   for preventing or treating the disease resulting from intracellular
CC   protozoan parasite infection such as malaria, tuberculosis,
CC   toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC   achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC   Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC   encephalitis, influenza, rubeola and dengue viral infections,
CC   poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC   respiratory syndrome. This sequence represent a nucleic acid used in the
CC   method of the invention.
XX
SQ   Sequence 228 BP; 63 A; 58 C; 62 G; 45 T; 0 U; 0 Other;

  Query Match             10.4%;  Score 228;  DB 13;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 60

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 120

Qy       1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1436
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 228


RESULT 85
ADR12558
ID   ADR12558 standard; DNA; 228 BP.
XX
AC   ADR12558;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Gene vaccine nucleic acid #13.
XX
KW   ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW   antibacterial; haemostatic; protozoacide; antiinflammatory;
KW   neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW   intracellular parasite; protozoan parasite infection; malaria;
KW   tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW   cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW   Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW   Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW   Herpes virus; severe acute respiratory syndrome.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .228
FT                   /*tag=  a
```

XX
PN WO2004067040-A1.
XX
PD 12-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-JP000975.
XX
PR 31-JAN-2003; 2003JP-00023507.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12559.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 25; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 228 BP; 64 A; 60 C; 60 G; 44 T; 0 U; 0 Other;

Query Match 10.4%; Score 228; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       981 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGT 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGT 60

Qy      1041 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy      1101 CAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAAC 1160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAAC 180

Qy      1161 ATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 228
```

RESULT 86
ADR12564
ID ADR12564 standard; DNA; 228 BP.
XX
AC ADR12564;
XX
DT 04-NOV-2004 (first entry)
XX
DE Gene vaccine nucleic acid #16.
XX
KW ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW antibacterial; haemostatic; protozoacide; antiinflammatory;
KW neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW intracellular parasite; protozoan parasite infection; malaria;
KW tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW Herpes virus; severe acute respiratory syndrome.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .228
FT /*tag= a
XX
PN WO2004067040-A1.
XX
PD 12-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-JP000975.
XX
PR 31-JAN-2003; 2003JP-00023507.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12565.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 31; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 228 BP; 64 A; 60 C; 61 G; 43 T; 0 U; 0 Other;

Query Match 10.4%; Score 228; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy      1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy      1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy      1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1892
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 228
```

RESULT 87
ADS87013
ID ADS87013 standard; DNA; 228 BP.
XX
AC ADS87013;

```
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 29.
XX
KW   vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW   Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW   lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW   colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS   Homo sapiens.
XX
PN   WO2004035085-A1.
XX
PD   29-APR-2004.
XX
PF   16-OCT-2003; 2003WO-JP013279.
XX
PR   17-OCT-2002; 2002JP-00302816.
XX
PA   (KYUS-) KYUSHU TLO CO LTD.
XX
PI   Himeno K,  Furue M,  Maehara Y;
XX
DR   WPI; 2004-357144/33.
DR   P-PSDB; ADS87014.
XX
PT   Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT   or cytokine genes for prevention and treatment of cancer.
XX
PS   Disclosure; SEQ ID NO 29; 266pp; Japanese.
XX
CC   The invention relates to a novel genetic vaccine containing the ubiquitin
CC   gene together with a gene encoding an antigenic protein containing a T-
CC   cell target sequence. The vaccine of the invention may be useful for
CC   prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC   (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC   of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC   colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC   is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC   invention.
XX
SQ   Sequence 228 BP; 63 A; 58 C; 62 G; 45 T; 0 U; 0 Other;

  Query Match             10.4%;  Score 228;  DB 13;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 60

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 120

Qy       1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1436
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 228


RESULT 88
ADS87017
ID   ADS87017 standard; DNA; 228 BP.
XX
AC   ADS87017;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 33.
XX
KW   vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW   Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW   lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW   colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS   Homo sapiens.
XX
PN   WO2004035085-A1.
XX
PD   29-APR-2004.
XX
PF   16-OCT-2003; 2003WO-JP013279.
XX
PR   17-OCT-2002; 2002JP-00302816.
XX
PA   (KYUS-) KYUSHU TLO CO LTD.
XX
PI   Himeno K,  Furue M,  Maehara Y;
XX
DR   WPI; 2004-357144/33.
DR   P-PSDB; ADS87018.
XX
PT   Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT   or cytokine genes for prevention and treatment of cancer.
XX
PS   Disclosure; SEQ ID NO 33; 266pp; Japanese.
XX
CC   The invention relates to a novel genetic vaccine containing the ubiquitin
CC   gene together with a gene encoding an antigenic protein containing a T-
CC   cell target sequence. The vaccine of the invention may be useful for
CC   prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC   (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC   of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC   colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC   is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC   invention.
XX
SQ   Sequence 228 BP; 64 A; 60 C; 61 G; 43 T; 0 U; 0 Other;

  Query Match             10.4%;  Score 228;  DB 13;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-102;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1892
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 228


RESULT 89
ADS87011
ID   ADS87011 standard; DNA; 228 BP.
XX
AC   ADS87011;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 27.
XX
KW   vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
```

KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
DR P-PSDB; ADS87012.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Disclosure; SEQ ID NO 27; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC invention.
XX
SQ Sequence 228 BP; 64 A; 60 C; 60 G; 44 T; 0 U; 0 Other;

Query Match 10.4%; Score 228; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-102;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 981 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGT 1040
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGT 60

Qy 1041 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1100
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy 1101 CAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAAC 1160
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAAC 180

Qy 1161 ATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 ATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 228

RESULT 90
ABT07115
ID ABT07115 standard; cDNA; 440 BP.
XX
AC ABT07115;
XX
DT 07-NOV-2002 (first entry)
XX
DE Human ovarian cancer associated coding sequence SEQ ID NO: 277.
XX
KW Human; ovarian cancer; cancer; gene; ss.
XX
OS Homo sapiens.
XX
PN US2002076715-A1.
XX
PD 20-JUN-2002.
XX
PF 06-JUN-2001; 2001US-00876889.
XX
PR 23-SEP-1998; 98US-00159320.
PR 08-FEB-1999; 99US-00246429.
PR 16-SEP-1999; 99US-00397787.
XX
PA (BENS/) BENSON D R.
PA (LODE/) LODES M J.
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
XX
PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX
DR WPI; 2002-598720/64.
XX
PT Composition for detecting and treating ovarian cancer, comprises a
PT specific polypeptide, polynucleotide, T cell population, or antigen
PT presenting cell.
XX
PS Example 1; Page 104; 188pp; English.
XX
CC The present invention relates to a method of detecting the presence of
CC ovarian cancer in a patient, involving detecting ovarian cancer
CC associated polynucleotides. The method is not only used to detect the
CC presence of cancer, preferably ovarian cancer in a patient, but also is
CC used to stimulate and/or expand T cells specific for an ovarian tumour
CC protein. The sequences can be used in vaccines used to treat cancer. The
CC present sequence is an ovarian cancer associated coding sequence
XX
SQ Sequence 440 BP; 125 A; 124 C; 107 G; 80 T; 0 U; 4 Other;

Query Match 9.6%; Score 211; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.2e-94;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 60

Qy 816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy 876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy 936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 966
||||||||||||||||||||||||||||||
Db 181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 211

RESULT 91
ABX72993
ID ABX72993 standard; cDNA; 440 BP.
XX
AC ABX72993;
XX
DT 14-MAR-2003 (first entry)
XX
DE Human ovarian carcinoma antigen partial cDNA sequence #259.
XX
KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17; OV18;
KW OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX
OS Homo sapiens.
XX
PN US6468758-B1.

XX
PD 22-OCT-2002.
XX
PF 16-SEP-1999; 99US-00397787.
XX
PR 23-SEP-1998; 98US-00159320.
PR 08-FEB-1999; 99US-00246429.
XX
PA (CORI-) CORIXA CORP.
XX
PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX
DR WPI; 2003-147101/14.
XX
PT Determining presence or absence of cancer in patient by contacting
PT patient sample with oligonucleotide that hybridizes to polynucleotide
PT encoding ovarian carcinoma antigen, and detecting amount of hybridization
PT in sample.
XX
PS Example 1; Col 201-202; 152pp; English.
XX
CC The present invention relates to compositions and methods for determining
CC presence or absence of cancer in patient. The method comprising
CC contacting a biological sample with an oligonucleotide that hybridises to
CC partial polynucleotide sequence encoding for human ovarian carcinoma
CC antigen, or its complement. The method and compositions are useful for
CC the therapy and diagnosis of ovarian cancer. The method can be used to
CC identify tumour antigens that are secreted from ovarian carcinoma and/or
CC other tumours. Effective cancer detection is achieved using the method of
CC the invention. ABX72720-ABX73050 represent partial cDNA sequences
CC encoding human ovarian carcinoma antigens
XX
SQ Sequence 440 BP; 125 A; 124 C; 107 G; 80 T; 0 U; 4 Other;

Query Match 9.6%; Score 211; DB 8; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.2e-94;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 60

Qy 816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy 876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy 936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 966
|||||||||||||||||||||||||||||||
Db 181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 211

RESULT 92
ADR12562
ID ADR12562 standard; DNA; 228 BP.
XX
AC ADR12562;
XX
DT 04-NOV-2004 (first entry)
XX
DE Gene vaccine nucleic acid #15.
XX
KW ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW antibacterial; haemostatic; protozoacide; antiinflammatory;
KW neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW intracellular parasite; protozoan parasite infection; malaria;
KW tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW Herpes virus; severe acute respiratory syndrome.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .228
FT /*tag= a
XX
PN WO2004067040-A1.
XX
PD 12-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-JP000975.
XX
PR 31-JAN-2003; 2003JP-00023507.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12563.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 29; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 228 BP; 64 A; 59 C; 60 G; 45 T; 0 U; 0 Other;

Query Match 9.5%; Score 209; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.2e-93;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy 1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy 1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy 1617 ATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
|||||||||||||||||||||||||||||
Db 181 ATCCAGAAAGAGTCCACCCTGCACCTGGT 209

RESULT 93
ADS87015
ID ADS87015 standard; DNA; 228 BP.
XX
AC ADS87015;
XX
DT 18-NOV-2004 (first entry)
XX

DE Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 31.
XX
KW vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
DR P-PSDB; ADS87016.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Disclosure; SEQ ID NO 31; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC invention.
XX
SQ Sequence 228 BP; 64 A; 59 C; 60 G; 45 T; 0 U; 0 Other;

Query Match 9.5%; Score 209; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.2e-93;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1617 ATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             |||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGT 209
```

RESULT 94
AAH48763
ID AAH48763 standard; DNA; 7648 BP.
XX
AC AAH48763;
XX
DT 29-OCT-2001 (first entry)
XX
DE Vector pHA-Ub DNA.
XX
KW Ubiquitin; protease function; protein metabolism; toxicology;
KW cell biology; vector; ds.
XX
OS Synthetic.
XX
PN DE10064343-A1.
XX
PD 16-AUG-2001.
XX
PF 21-DEC-2000; 2000DE-01064343.
XX
PR 21-DEC-2000; 2000DE-01064343.
XX
PA (CULL/) CULLEN P.
PA (WENN/) WENNER C.
XX
PI Cullen P, Wenner C;
XX
DR WPI; 2001-497721/55.
XX
PT Investigate intracellular protein metabolism using new nucleic acid
PT construct, comprising multiple ubiquitin coding sections and a selection
PT marker.
XX
PS Example 1; Fig 7; 16pp; German.
XX
CC This invention describes a novel nucleic acid (I), comprising at least
CC one ubiquitin coding section and a selection marker fixed at a separate
CC location from the ubiquitin coding section. The invention is used to
CC investigate protease function in protein metabolism. This is useful in
CC toxicology and general cell biology. The invention provides a method that
CC is cheaper, quicker, and more reproducible than prior art methods where
CC the ubiquitin is only transiently expressed. This sequence represents the
CC vector pHA-Ub encoding DNA is used in the method of the invention
XX
SQ Sequence 7648 BP; 1850 A; 2036 C; 1947 G; 1815 T; 0 U; 0 Other;

Query Match 9.5%; Score 209; DB 4; Length 7648;
Best Local Similarity 100.0%; Pred. No. 3e-93;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1048

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1108

Qy       1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1168

Qy       1617 ATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             |||||||||||||||||||||||||||||
Db       1169 ATCCAGAAAGAGTCCACCCTGCACCTGGT 1197
```

RESULT 95
ABK69710
ID ABK69710 standard; cDNA; 228 BP.
XX
AC ABK69710;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human ubiquitin open reading frame.
XX
KW Human; mouse; cytostatic; immunostimulant; WT1; cancer; immune response;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200228414-A1.

```
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US031139.
XX
PR   06-OCT-2000; 2000US-00684361.
PR   09-OCT-2000; 2000US-00685830.
PR   15-FEB-2001; 2001US-00785019.
PR   24-AUG-2001; 2001US-00938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A, Mcneill PD, Smithgall M, Moulton G, Vedvick TS;
PI   Sleath PR, Mossman S, Evans L, Spies AG, Boydston J;
XX
DR   WPI; 2002-352217/38.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient.
XX
PS   Claim 1; Page 239; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides are
CC   used for treating and detecting cancer in a patient, and for stimulating
CC   an immune response in patient. ABK69648-ABK69724 represent WT1 coding
CC   sequences and PCR primers of the invention
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 96
ADB67586
ID   ADB67586 standard; cDNA; 228 BP.
XX
AC   ADB67586;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human ubiquitin cDNA.
XX
KW   Human; Wilm's tumour antigen; WT1; cytostatic; antigen; T cell;
KW   antigen-presenting cell; CD4+; CD8+; immune response; cancer; vaccine;
KW   leukaemia; antigen; ss; gene.
XX
OS   Homo sapiens.
XX
PN   US2003072767-A1.
XX
PD   17-APR-2003.
XX
PF   24-AUG-2001; 2001US-00938864.
XX
PR   30-SEP-1998;   98US-00164223.
PR   25-MAR-1999;   99US-00276484.
PR   06-OCT-2000; 2000US-00684361.
PR   09-OCT-2000; 2000US-00685830.
PR   15-FEB-2001; 2001US-00785019.
XX
PA   (GAIG/) GAIGER A.
PA   (MCNE/) MCNEILL P D.
PA   (SMIT/) SMITHGALL M.
PA   (MOUL/) MOULTON G.
PA   (VEDV/) VEDVICK T S.
PA   (SLEA/) SLEATH P R.
PA   (MOSS/) MOSSMAN S.
PA   (EVAN/) EVANS L.
PA   (SPIE/) SPIES A G.
PA   (BOYD/) BOYDSTON J.
XX
PI   Gaiger A, Mcneill PD, Smithgall M, Moulton G, Vedvick TS;
PI   Sleath PR, Mossman S, Evans L, Spies AG, Boydston J;
XX
DR   WPI; 2003-605957/57.
XX
PT   Novel Wilms tumor polynucleotides encoding WT1 polypeptides, useful for
PT   detecting the presence of cancer in a patient, and in pharmaceutical
PT   compositions, e.g. vaccines, for treating cancers e.g. leukemia.
XX
PS   Claim 1; Page 143; 197pp; English.
XX
CC   The invention relates to an isolated Wilm tumour (WT1) polynucleotide
CC   comprising a full length protein, truncated protein, mutated protein or
CC   fusion protein. Also included are the encoded WT1 proteins, expression
CC   vectors, host cells, antibodies, detecting the presence of a cancer in a
CC   patient by contacting a biological sample with a binding agent that binds
CC   to a WT1 protein, an oligonucleotide that hybridises to a WT1
CC   polynucleotide, stimulating and/or expanding T cells specific for a
CC   tumour protein by contacting T cells with the WT1 polynucleotide, the WT1
CC   protein or antigen-presenting cells that express the WT1 protein, a
CC   composition (C1) (comprising physiologically acceptable carriers and
CC   immunostimulants as first component, and a second component selected from
CC   the WT1 polynucleotide, the WT1 protein, Ab, or T cells and antigen
CC   presenting cells that express the WT1 protein), inhibiting the
CC   development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
CC   cells isolated from a patient with the WT1 polynucleotide, the WT1
CC   protein or antigen presenting cells that express the WT1 protein, such
CC   that T cell proliferate, administering the proliferated T cells to the
CC   patient, and thus inhibiting the development of a cancer in the patient)
CC   and a composition (C2) (comprising a WT1 polypeptide resuspended in a
CC   buffer comprising at least 1-3 sugars selected from trehalose, maltose,
CC   sucrose, fructose and glucose, at a concentration of 7-13%, and
CC   optionally ethanolamine, cysteine and Polysorbate-80, or WT1 polypeptide
CC   and MPL-SE or Enhanzyn). Also disclosed as anew are polypeptides
CC   comprising a variant of an immunogenic portion of WT1 polypeptide. C1 is
CC   useful for stimulating immune response in a patient, and for treating
CC   cancer in a patient. The oligonucleotide is also useful for determining
CC   the presence of a cancer in a patient. The WT1 polynucleotide and the WT1
CC   protein are useful in pharmaceutical compositions, e.g. vaccines. the WT1
CC   protein is useful as marker to indicate the presence or absence of a
CC   cancer. C1 is useful for inhibiting the development of a malignant
CC   disease in a patient, for preventing and treating metastatic diseases
CC   e.g. leukaemia and cancer, and for removing tumour cells from a
CC   biological sample. Ab (binding agent for the WT1 protein) is useful for
CC   detecting the presence of cancer in a patient. The present sequence
CC   encodes a fusion partner for WT1.
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 10;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 97
ADG75008
ID   ADG75008 standard; DNA; 228 BP.
XX
AC   ADG75008;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human ubiquitin DNA - SEQ ID 80.
XX
KW   immunogenic; herpes simplex virus type 2; HSV; virucide; HSV2 infection;
KW   vaccine; ds; ubiquitin; human.
XX
OS   Homo sapiens.
XX
PN   WO2003086308-A2.
XX
PD   23-OCT-2003.
XX
PF   09-APR-2003; 2003WO-US011231.
XX
PR   11-APR-2002; 2002US-00121988.
PR   19-JUL-2002; 2002US-00200562.
PR   06-SEP-2002; 2002US-00237551.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Day CH,  Hosken NA,  Parsons JM;
XX
DR   WPI; 2003-845255/78.
XX
PT   New isolated polypeptide comprising at least an immunogenic portion of a
PT   herpes simplex virus (HSV) antigen, useful for diagnosing, treating or
PT   preventing HSV infection.
XX
PS   Claim 3; SEQ ID NO 80; 383pp; English.
XX
CC   The invention relates to a novel isolated polypeptide comprising at least
CC   an immunogenic portion of a herpes simplex virus (HSV) antigen. The
CC   polypeptide of the invention demonstrates virucide activity and may be
CC   useful for diagnosing, treating or preventing Human herpesvirus 2 (herpes
CC   simplex virus type 2, HSV2) infection via the production of a vaccine.
CC   The current sequence is that of an human ubiquitin DNA which was used in
CC   the exemplification of the invention.
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 10;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141,

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 98
ADJ80819
ID   ADJ80819 standard; DNA; 228 BP.
XX
AC   ADJ80819;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Wilm's tumor antigen associated DNA #18.
XX
KW   ds; gene; cytostatic; gene therapy; vaccine; Wilm's tumor antigen; WT1;
KW   cancer; leukemia.
XX
OS   Homo sapiens.
XX
PN   WO2003037060-A2.
XX
PD   08-MAY-2003.
XX
PF   30-OCT-2002; 2002WO-US035194.
XX
PR   30-OCT-2001; 2001US-00002603.
PR   16-APR-2002; 2002US-00125635.
PR   12-JUL-2002; 2002US-00195835.
PR   16-SEP-2002; 2002US-00244830.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Mcneill PD,  Jaya N,  Carter D;
XX
DR   WPI; 2003-430453/40.
XX
PT   New isolated polypeptide for diagnosing, preventing or treating malignant
PT   diseases, e.g. cancer or leukemia, comprises an immunogenic portion of a
PT   Wilms' tumor antigen.
XX
PS   Disclosure; SEQ ID NO 384; 371pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising an
CC   immunogenic portion of a Wilm's tumor antigen, or its variant that
CC   differs in one or more substitutions, deletions, additions and/or
CC   insertions, such that the ability of the variant to react with WT1-
CC   specific antisera and/or T-cell lines or clones is not substantially
CC   diminished. The composition and methods are useful in diagnosing,
CC   preventing or treating malignant diseases, such as cancer or leukemia.
CC   This sequence represents a DNA of the invention.
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 10;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
                ||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 99
ADJ83739
ID   ADJ83739 standard; DNA; 228 BP.
XX
AC   ADJ83739;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human ubiquitin ORF DNA.
XX
KW   immune response; WT1; Wilms tumour; immunostimulant; cytostatic; vaccine;
KW   malignant disease; leukaemia; cancer; human; ds; ubiquitin.
XX
OS   Homo sapiens.
XX
PN   US2003235557-A1.
XX
PD   25-DEC-2003.
XX
PF   16-SEP-2002; 2002US-00244830.
XX
PR   30-SEP-1998;   98US-00164223.
PR   25-MAR-1999;   99US-00276484.
PR   06-OCT-2000; 2000US-00684361.
PR   09-OCT-2000; 2000US-00685830.
PR   15-FEB-2001; 2001US-00785019.
PR   24-AUG-2001; 2001US-00938864.
PR   30-OCT-2001; 2001US-00002603.
PR   16-APR-2002; 2002US-00125635.
PR   12-JUL-2002; 2002US-00195835.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Gaiger A,  Cheever MA;
XX
DR   WPI; 2004-070542/07.
XX
PT   Inducing an immune response in an animal, useful for treating or
PT   preventing cancer, comprises administering a composition comprising a
PT   viral vector comprising an immunogenic portion of a WT1 polynucleotide.
XX
PS   Example 19; SEQ ID NO 384; 229pp; English.
XX
CC   The invention relates to a novel method for inducing an immune response
CC   in an animal which involves administering a first and a second
CC   composition comprising a first and a second viral vector, respectively,
CC   each having at least an immunogenic portion of a WT1 (Wilms tumour)
CC   polynucleotide operably linked to an expression control sequence. The
CC   method of the invention has immunostimulant and cytostatic applications
CC   and may be useful for generating or enhancing an immune response to WT1,
CC   possibly via the production of a vaccine, as well as for preventing
CC   and/or treating malignant diseases such as leukaemia and cancer. The
CC   current sequence is that of the human WT1-related DNA of the invention.
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 12;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
                ||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 100
ADL57627
ID   ADL57627 standard; cDNA; 228 BP.
XX
AC   ADL57627;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human ubiquitin open reading frame.
XX
KW   Human; Wilms' tumour; WT1; cancer; cytostatic; immunogen; antigen;
KW   vaccine; T-cell; MHC; Major Histocompatibility Complex; HLA-A2;
KW   human leukocyte antigen; Ra12; HIS tag; twin arginine translocator; TAT;
KW   malignant disease; ss; gene.
XX
OS   Homo sapiens.
XX
PN   US2003215458-A1.
XX
PD   20-NOV-2003.
XX
PF   30-OCT-2002; 2002US-00286333.
XX
PR   30-SEP-1998;   98US-00164223.
PR   25-MAR-1999;   99US-00276484.
PR   06-OCT-2000; 2000US-00684361.
PR   09-OCT-2000; 2000US-00685830.
PR   15-FEB-2001; 2001US-00785019.
PR   24-AUG-2001; 2001US-00938864.
PR   30-OCT-2001; 2001US-00002603.
PR   16-APR-2002; 2002US-00125635.
PR   12-JUL-2002; 2002US-00195835.
PR   16-SEP-2002; 2002US-00244830.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Gaiger A,  Mcneill PD,  Jaya N;
XX
DR   WPI; 2004-021827/02.
XX
PT   New polypeptides and polynucleotides useful for generating or enhancing
PT   an immune response to Wilms' tumor, or for treating and/or preventing
PT   malignant diseases, e.g. cancer.
XX
PS   Example 19; SEQ ID NO 384; 259pp; English.
XX
CC   The invention relates to an isolated polypeptide comprising an
CC   immunogenic portion of a Wilms' tumour antigen (WT1), or their variants
CC   that differs in substitution(s), deletion(s), addition(s) and/or
CC   insertion(s), where ability of the variant to react with WT1-specific
CC   antisera and/or T-cell lines or clones is not substantially diminished.
CC   Also included are a fusion protein comprising at least one WT1
CC   polypeptide, an isolated polynucleotide encoding the fusion protein, a
CC   composition comprising a WT1 polypeptide in combination with a
CC   pharmaceutical carrier or excipient, a vaccine comprising a WT1
CC   polypeptide in combination with a non-specific immune response enhancer,
CC   an expression vector comprising the polynucleotide operably linked to an
CC   expression control sequence and a host cell transformed or transfected
CC   with the expression vector. The immunogenic portion of the WT1
CC   polypeptide has been modified such that the ability of the immunogenic
CC   portion to bind to an MHC (Major Histocompatibility Complex) molecule or
```

```
CC   to HLA-A2 is increased relative to that of the immunogenic portion. The
CC   WT1 polypeptide may comprise a Wilms' tumour antigen having a deletion of
CC   a proline rich region at amino acid positions 54-68 of the Wilms' tumour
CC   antigen. The fusion partner is selected from Ra12, protein D, LYTA, a HIS
CC   tag, a targeting signal capable of directing a polypeptide to the
CC   endosomal/lysosomal compartment, twin arginine translocator (TAT), and
CC   truncated twin arginine translocator. The polypeptide, polynucleotide or
CC   compositions comprising them may be used for generating or enhancing an
CC   immune response to Wilms' tumour, and for treating and/or preventing
CC   malignant diseases, e.g. cancer. The compositions may be used as markers
CC   for the progression of cancer. The present sequence is a human cDNA for a
CC   fusion partner for WT-1.
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 12;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 101
ADO09182
ID   ADO09182 standard; DNA; 228 BP.
XX
AC   ADO09182;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human ubiquitin DNA.
XX
KW   vaccine; malignant disease; immunogenic; leukaemia; cancer; ds; gene;
KW   human; ubiquitin.
XX
OS   Homo sapiens.
XX
PN   US2004018204-A1.
XX
PD   29-JAN-2004.
XX
PF   30-APR-2003; 2003US-00427717.
XX
PR   24-AUG-2001; 2001US-00938864.
PR   30-OCT-2001; 2001US-00002603.
PR   16-APR-2002; 2002US-00125635.
PR   12-JUL-2002; 2002US-00195835.
PR   16-SEP-2002; 2002US-00244830.
PR   30-OCT-2002; 2002US-00286333.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Gaiger A,  Mcneill PD,  Jaya N;
XX
DR   WPI; 2004-121992/12.
XX
PT   New polypeptides and polynucleotides comprising an immunogenic portion of
PT   a Wilms' tumor antigen, useful as vaccines for preventing or treating
PT   malignant diseases, e.g. cancer or leukemia.
XX
PS   Example 19; SEQ ID NO 384; 252pp; English.
XX
CC   The invention relates to an isolated polypeptide which comprises an
CC   immunogenic portion of a Wilms' tumour antigen having a deletion of a
CC   proline rich region, so that the ability of the variant to react with WT1
CC   -specific antisera and/or T-cell lines or clones is not substantially
CC   diminished. The polypeptide is useful as a vaccine in combination with a
CC   non-specific immune response enhancer for preventing malignant diseases.
CC   The polypeptide is useful for enhancing or inducing an immune response in
CC   a patient; for inhibiting the development of malignant disease in a
CC   patient; for stimulating or expanding T cells. The polypeptides,
CC   polynucleotides and compositions are useful for treating malignant
CC   diseases including leukaemia and cancer. The present sequence represents
CC   the human ubiquitin DNA.
XX
SQ   Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

  Query Match             9.4%;  Score 207;  DB 12;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 102
ADR12534
ID   ADR12534 standard; DNA; 228 BP.
XX
AC   ADR12534;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Gene vaccine nucleic acid #1.
XX
KW   ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW   antibacterial; haemostatic; protozoacide; antiinflammatory;
KW   neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW   intracellular parasite; protozoan parasite infection; malaria;
KW   tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW   cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW   Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW   Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW   Herpes virus; severe acute respiratory syndrome.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .228
FT                   /*tag=  a
XX
PN   WO2004067040-A1.
XX
PD   12-AUG-2004.
XX
PF   30-JAN-2004; 2004WO-JP000975.
XX
PR   31-JAN-2003; 2003JP-00023507.
XX
```

PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12535.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 1; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

Query Match 9.4%; Score 207; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.2e-92;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228
```

RESULT 103
ADS86985
ID ADS86985 standard; DNA; 228 BP.
XX
AC ADS86985;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genetic vaccine-related ubiquitin (Ub) DNA.
XX
KW vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; ds; gene; human.
XX
OS Homo sapiens.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
DR P-PSDB; ADS86986.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Disclosure; SEQ ID NO 1; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a genetic vaccine-related ubiquitin (Ub) DNA of the invention.
XX
SQ Sequence 228 BP; 64 A; 58 C; 59 G; 47 T; 0 U; 0 Other;

Query Match 9.4%; Score 207; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.2e-92;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228
```

RESULT 104
AAL51106
ID AAL51106 standard; DNA; 231 BP.
XX
AC AAL51106;
XX
DT 27-FEB-2003 (first entry)
XX
DE Human ubiquitin coding sequence.
XX
KW Gene; ds; gene therapy; vaccine; humoral immune response;
KW cellular immune response; immune response modulation;
KW pathogenic infection; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200283181-A1.
XX
PD 24-OCT-2002.
XX
PF 18-APR-2002; 2002WO-AU000486.
XX
PR 18-APR-2001; 2001AU-00004468.
XX
PA (UYQU) UNIV QUEENSLAND.
XX
PI Frazer IH;
XX

DR WPI; 2003-075509/07.
DR P-PSDB; AAO16088.
XX
PT New compositions having antigens, polynucleotides encoding the antigens,
PT or antigen-presenting cells, useful for modulating an immune response,
PT e.g. for treating or preventing pathogenic infections or rheumatoid
PT arthritis.
XX
PS Disclosure; Page 84; 139pp; English.
XX
CC The invention comprises compositions for eliciting a humoral or cellular
CC immune response against a target antigen. The compositions of the
CC invention are useful for eliciting a humoral and cellular immune response
CC against a target antigen, modulating an immune response in a patient, and
CC the treatment or prophylaxis of a disease or condition. This includes
CC pathogenic infections (e.g. viral, bacterial, fungal or protozoan), or an
CC immune response to an autoantigen (e.g. rheumatoid arthritis). The
CC present DNA sequence represents a gene which was used in the invention
XX
SQ Sequence 231 BP; 65 A; 58 C; 60 G; 48 T; 0 U; 0 Other;

Query Match 9.4%; Score 207; DB 8; Length 231;
Best Local Similarity 100.0%; Pred. No. 3.2e-92;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy 606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy 666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy 726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
|||||||||||||||||||||||||||
Db 202 CACCTGGTACTCCGTCTCAGAGGTGGG 228

RESULT 105
AAA16164/c
ID AAA16164 standard; DNA; 609 BP.
XX
AC AAA16164;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human colon cancer differentially expressed nucleotide sequence #169.
XX
KW Colon cancer; detect; differential expression; human; treatment;
KW detect mutation; non-invasive diagnostic method; ds.
XX
OS Homo sapiens.
XX
PN WO200012702-A2.
XX
PD 09-MAR-2000.
XX
PF 30-AUG-1999; 99WO-US019424.
XX
PR 31-AUG-1998; 98US-0098639P.
PR 27-JAN-1999; 99US-0117393P.
XX
PA (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Carroll E;
PI Catino TJ, Dwivedi P, Ford DM, Lewis ME, Molino GA, Monahan JE;
PI Schlegel R;
XX
DR WPI; 2000-256641/22.
XX
PT Novel nucleic acids and proteins for identifying therapeutic agents
PT useful for treating and diagnosing cancer, especially colon cancer.
XX
PS Claim 16; Page 191-192; 345pp; English.
XX
CC This sequence represents a human nucleotide sequence which is
CC differentially expressed in colon cancer cells compared to the expression
CC levels in normal cells. The nucleotide sequence can be used as a source
CC of primers and probes. The nucleotide sequence is useful for determining
CC the phenotype of a cell by detecting the differential expression of the
CC sequence relative to a normal cell. The probes derived from the sequence
CC can also be used to determine the phenotype of cells in a sample. Probes
CC and antibodies which hybridise to the nucleotide sequence can also be
CC used to determine the phenotype of a cell. The primers are useful for
CC detecting a mutation in a test nucleotide sequence and also for detecting
CC cancer, preferably colon cancer. Antibodies against the protein encoded
CC by the nucleotide sequence can also be used in a method to detect colon
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing
CC colon cancer at an early stage
XX
SQ Sequence 609 BP; 123 A; 160 C; 138 G; 174 T; 0 U; 14 Other;

Query Match 9.4%; Score 207; DB 3; Length 609;
Best Local Similarity 100.0%; Pred. No. 3.1e-92;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 364 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 305

Qy 1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 304 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 245

Qy 1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 244 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 185

Qy 1871 GGTGCTCCGTCTCAGAGGTGGGATGCA 1897
|||||||||||||||||||||||||||
Db 184 GGTGCTCCGTCTCAGAGGTGGGATGCA 158

RESULT 106
ADR23183
ID ADR23183 standard; DNA; 1581 BP.
XX
AC ADR23183;
XX
DT 04-NOV-2004 (first entry)
XX
DE Yellow fluorescent protein-beta-arrestin 2-ubiquitin coding sequence.
XX
KW Arrestin; beta-arrestin 2; yellow fluorescent protein; ubiquitin;
KW G-protein coupled receptor; gene therapy; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .1581
FT /*tag= a
FT /product= "YEFP-beta arrestin 2-ubiquitin"
FT /partial
FT /note= "No start codon"
XX
PN WO2004067717-A2.
XX
PD 12-AUG-2004.
XX
PF 26-JAN-2004; 2004WO-US002029.
XX
PR 24-JAN-2003; 2003US-0442403P.

XX
PA (UYDU-) UNIV DUKE.
XX
PI Shenoy S, Lefkowitz RJ;
XX
DR WPI; 2004-580992/56.
DR P-PSDB; ADR23186.
XX
PT New modified arrestin comprising an arrestin and a ubiquitin moiety or
PT their biologically active fragment, useful for treating a human or a non-
PT human subject suffering from a GPCR-related disease.
XX
PS Claim 9; Fig 8; 90pp; English.
XX
CC The present sequence is the coding sequence for a chimeric polypeptide
CC comprising enhanced yellow fluorescent protein (EYFP), beta-arrestin 2
CC and ubiquitin. The chimeric polypeptide is a claimed example of modified
CC arrestins of the invention that comprise an arrestin (or biologically
CC active fragment) and ubiquitin (or biologically active fragment) and an
CC optional label. The modified arrestin has an increased affinity for a G-
CC protein coupled receptor (GPCR) such that the arrestin remains associated
CC with the GPCR and traffics with the GPCR into endosomes and hence does
CC not dissociate at or near the plasma membrane. The modified arrestin
CC increases the internalisation of the GPCR. The invention provides the
CC modified arrestins, nucleic acids encoding them, expression vectors and
CC host (mammalian, bacterial, yeast, plant, insect or animal) cells. The
CC modified arrestins are used in a claimed method of screening compounds
CC and sample solutions for a GPCR agonist, antagonist, inverse agonist or
CC desensitization active compound. Claimed methods of treating a human or
CC non-human subject suffering from a GPCR-related disease comprise
CC administering a compound identified by this method or a nucleic acid
CC encoding a modified arrestin.
XX
SQ Sequence 1581 BP; 387 A; 467 C; 401 G; 326 T; 0 U; 0 Other;

Query Match 9.4%; Score 207; DB 13; Length 1581;
Best Local Similarity 100.0%; Pred. No. 3e-92;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 1431

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 1491

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 1551

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              ||||||||||||||||||||||||||
Db       1552 CACCTGGTACTCCGTCTCAGAGGTGGG 1578
```

RESULT 107
AAC98259
ID AAC98259 standard; cDNA; 332 BP.
XX
AC AAC98259;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:269.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200055351-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005883.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587534/55.
DR P-PSDB; AAB53502.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.
XX
PS Claim 1; Page 694; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 332 BP; 84 A; 84 C; 87 G; 71 T; 0 U; 6 Other;

Query Match 9.4%; Score 205; DB 3; Length 332;
Best Local Similarity 100.0%; Pred. No. 3.1e-91;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 94

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 154

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 214

Qy        229 GCACCCTGTCTGACTACAACATCCA 253
              |||||||||||||||||||||||||
Db        215 GCACCCTGTCTGACTACAACATCCA 239
```

RESULT 108
ACF04785
ID ACF04785 standard; DNA; 480 BP.
XX
AC ACF04785;
XX
DT 18-DEC-2003 (first entry)
XX
DE Vector pG31010 containing toxoid antibody light chain Fv domain DNA.
XX

KW Antibody Fv chain; stabilisation; assembly; alpha-helical coiled-coil;
KW AHEC; antiparallel; drug targeting; vector; gene; ds.
XX
OS Synthetic.
XX
PN WO2003066660-A2.
XX
PD 14-AUG-2003.
XX
PF 05-FEB-2003; 2003WO-EP001217.
XX
PR 05-FEB-2002; 2002US-0354376P.
XX
PA (IMMU-) IMMUNOLEX THERAPEUTICS APS.
XX
PI Caterer NR, Uttental LO, Neilson RW;
XX
DR WPI; 2003-679535/64.
DR P-PSDB; ABR84725.
XX
PT Composition useful for forming therapeutic antibodies and antibody
PT fragments comprises pair of antibody Fv fragments linked and stabilized
PT by antiparallel heterogeneous alpha-helical coiled-coil peptides.
XX
PS Example 2; Fig 7; 35pp; English.
XX
CC The present invention relates to a composition which comprises a pair of
CC antibody Fv fragments linked and stabilised by antiparallel heterogeneous
CC alpha-helical coiled-coil (AHEC) peptides. The composition is used to
CC form multimeric complexes and therapeutic antibodies and antibody
CC fragments useful for e.g. inhibition of receptor binding and the
CC targeting of drugs, toxins and labels in research, industry and
CC healthcare. The present sequence is a coding sequence of a vector
CC encoding an AHEC peptide used in the exemplification of the invention
XX
SQ Sequence 480 BP; 127 A; 137 C; 129 G; 87 T; 0 U; 0 Other;

Query Match 9.3%; Score 203; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 3e-90;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 97 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 156

Qy 606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 157 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 216

Qy 666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 217 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 276

Qy 726 CACCTGGTACTCCGTCTCAGAGG 748
|||||||||||||||||||||||
Db 277 CACCTGGTACTCCGTCTCAGAGG 299

RESULT 109
ACF04786
ID ACF04786 standard; DNA; 504 BP.
XX
AC ACF04786;
XX
DT 18-DEC-2003 (first entry)
XX
DE Vector pG31027 containing toxoid antibody light chain Fv domain DNA.
XX
KW Antibody Fv chain; stabilisation; assembly; alpha-helical coiled-coil;
KW AHEC; antiparallel; drug targeting; vector; gene; ds.
XX
OS Synthetic.
XX
PN WO2003066660-A2.
XX
PD 14-AUG-2003.
XX
PF 05-FEB-2003; 2003WO-EP001217.
XX
PR 05-FEB-2002; 2002US-0354376P.
XX
PA (IMMU-) IMMUNOLEX THERAPEUTICS APS.
XX
PI Caterer NR, Uttental LO, Neilson RW;
XX
DR WPI; 2003-679535/64.
DR P-PSDB; ABR84726.
XX
PT Composition useful for forming therapeutic antibodies and antibody
PT fragments comprises pair of antibody Fv fragments linked and stabilized
PT by antiparallel heterogeneous alpha-helical coiled-coil peptides.
XX
PS Example 2; Fig 8; 35pp; English.
XX
CC The present invention relates to a composition which comprises a pair of
CC antibody Fv fragments linked and stabilised by antiparallel heterogeneous
CC alpha-helical coiled-coil (AHEC) peptides. The composition is used to
CC form multimeric complexes and therapeutic antibodies and antibody
CC fragments useful for e.g. inhibition of receptor binding and the
CC targeting of drugs, toxins and labels in research, industry and
CC healthcare. The present sequence is a coding sequence of a vector
CC encoding an AHEC peptide used in the exemplification of the invention
XX
SQ Sequence 504 BP; 135 A; 140 C; 139 G; 90 T; 0 U; 0 Other;

Query Match 9.3%; Score 203; DB 10; Length 504;
Best Local Similarity 100.0%; Pred. No. 3e-90;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 180

Qy 606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 240

Qy 666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 300

Qy 726 CACCTGGTACTCCGTCTCAGAGG 748
|||||||||||||||||||||||
Db 301 CACCTGGTACTCCGTCTCAGAGG 323

RESULT 110
ACH34906
ID ACH34906 standard; cDNA; 471 BP.
XX
AC ACH34906;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human endothelial cell cDNA #3039.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.

```
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 22118; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 471 BP; 125 A; 101 C; 104 G; 140 T; 0 U; 1 Other;

  Query Match             9.2%;  Score 201;  DB 9;  Length 471;
  Best Local Similarity   100.0%;  Pred. No. 3e-89;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 109

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 169

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 229

Qy        229 GCACCCTGTCTGACTACAACA 249
              |||||||||||||||||||||
Db        230 GCACCCTGTCTGACTACAACA 250


RESULT 111
ADU14717
ID   ADU14717 standard; DNA; 549 BP.
XX
AC   ADU14717;
XX
DT   27-JAN-2005  (first entry)
XX
DE   Solid tumour prognosis gene seqid 5156.
XX
KW   cytostatic; gene therapy; expression profile; solid tumour;
KW   peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
OS   Unidentified.
XX
PN   WO2004097052-A2.
XX
PD   11-NOV-2004.
XX
PF   29-APR-2004; 2004WO-US013587.
XX
PR   29-APR-2003; 2003US-0466067P.
PR   23-JAN-2004; 2004US-0538246P.
XX
PA   (AMHP ) WYETH.
PA   (STRA/) STRAHS A.
XX
PI   Strahs A,  Trepicchio WL,  Burczynski ME,  Twine NC,  Slonim DK;
PI   Immerman F,  Dorner AJ;
XX
DR   WPI; 2004-804779/79.
XX
PT   A method, useful for prognosing and treating solid tumor, comprises
PT   comparing an expression profile of a gene expressed in peripheral blood
PT   mononuclear cells to a reference expression profile of a gene.
XX
PS   Disclosure; Page; 111pp; English.
XX
CC   The invention describes a method comprising comparing an expression
CC   profile of at least one gene in a peripheral blood sample of a patient to
CC   at least one reference expression profile of the at least one gene, where
CC   the patient has a solid tumour, and each of the gene is differentially
CC   expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC   of patients as compared to PBMCs of a second class of patients, where
CC   both the first and second classes of patients have the solid tumour, and
CC   each of the first and second classes is a subcluster formed by an
CC   unsupervised clustering analysis of gene expression profiles in PBMCs of
CC   a population of patients who have the solid tumour, and where the
CC   majority of the first class of patients has a first clinical outcome, and
CC   the majority of the second class of patients has a second clinical
CC   outcome. Also described are: a system comprising (i) a memory or a
CC   storage medium including data that represent an expression profile of at
CC   least one gene in a peripheral blood sample of a patient who has a solid
CC   tumour, (ii) at least another storage medium including data that
CC   represent at least one reference expression profile of the gene, (iii) a
CC   program capable of comparing the expression profile to the reference
CC   expression profile, and (iv) a processor capable of executing the
CC   program, where expression levels of the gene in peripheral blood
CC   mononuclear cells of patients who have the solid tumour correlate with
CC   clinical outcomes of the patients; and a nucleic acid or protein array
CC   comprising concentrated probes for solid tumour prognosis genes, where
CC   each of the solid tumour prognosis genes is differentially expressed in
CC   PBMCs of a first class of patients as compared to PBMCs of a second class
CC   of patients, where both the first and second classes of patients have a
CC   solid tumour, and where the first class of patients has a first clinical
CC   outcome, and the second class of patients has a second clinical outcome.
CC   The method, system, and array are useful for prognosing and treating
CC   solid tumours. This sequence represents a solid tumour prognosis gene of
CC   the invention. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 549 BP; 137 A; 130 C; 132 G; 96 T; 0 U; 54 Other;

  Query Match             9.2%;  Score 201;  DB 13;  Length 549;
  Best Local Similarity   100.0%;  Pred. No. 3e-89;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 194

Qy       1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 254
```

```
Qy       1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 314

Qy       1871 GGTGCTCCGTCTCAGAGGTGG 1891
             |||||||||||||||||||||
Db        315 GGTGCTCCGTCTCAGAGGTGG 335
```

RESULT 112
ADR12570
ID ADR12570 standard; DNA; 234 BP.
XX
AC ADR12570;
XX
DT 04-NOV-2004 (first entry)
XX
DE Gene vaccine nucleic acid #19.
XX
KW ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW antibacterial; haemostatic; protozoacide; antiinflammatory;
KW neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW intracellular parasite; protozoan parasite infection; malaria;
KW tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW Herpes virus; severe acute respiratory syndrome.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .228
FT /*tag= a
XX
PN WO2004067040-A1.
XX
PD 12-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-JP000975.
XX
PR 31-JAN-2003; 2003JP-00023507.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12571.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 37; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 234 BP; 64 A; 58 C; 64 G; 48 T; 0 U; 0 Other;

Query Match 8.9%; Score 195; DB 13; Length 234;
Best Local Similarity 100.0%; Pred. No. 2.9e-86;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1932 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 99

Qy       1992 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 159

Qy       2052 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTG 2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTG 219

Qy       2112 AGGGGGGGTGTCTAA 2126
             ||||||||||||||||
Db        220 AGGGGGGGTGTCTAA 234
```

RESULT 113
ADS87023
ID ADS87023 standard; DNA; 234 BP.
XX
AC ADS87023;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 39.
XX
KW vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
DR P-PSDB; ADS87024.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Disclosure; SEQ ID NO 39; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC invention.
XX
SQ Sequence 234 BP; 64 A; 58 C; 64 G; 48 T; 0 U; 0 Other;

Query Match 8.9%; Score 195; DB 13; Length 234;
Best Local Similarity 100.0%; Pred. No. 2.9e-86;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1932 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 99

Qy       1992 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 159

Qy       2052 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTG 2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTG 219

Qy       2112 AGGGGGGGTGTCTAA 2126
             ||||||||||||||||
Db        220 AGGGGGGGTGTCTAA 234
```

RESULT 114
AAA69785/c
ID AAA69785 standard; cDNA; 605 BP.
XX
AC AAA69785;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide SEQ ID NO:95.
XX
KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200036107-A2.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US030270.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
DR WPI; 2000-431589/37.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer.
XX
PS Claim 18; Fig 15; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557 represent human
CC ovarian carcinoma polynucleotides and proteins used in the
CC exemplification of the present invention
XX
SQ Sequence 605 BP; 111 A; 144 C; 140 G; 159 T; 0 U; 51 Other;

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db         179 AAAGAGTCCAC 169
```

RESULT 115
ABN72679/c
ID ABN72679 standard; DNA; 605 BP.
XX
AC ABN72679;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ovarian carcinoma antigen polynucleotide #95.
XX
KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US022635.
XX
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2002-164781/21.
XX
PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
PS Example 1; Page 131; 408pp; English.
XX
CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents DNA related
CC to the invention
XX
SQ Sequence 605 BP; 111 A; 144 C; 140 G; 159 T; 0 U; 51 Other;

Query Match 8.7%; Score 191; DB 6; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
                ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 116
ADA08844/c
ID ADA08844 standard; DNA; 605 BP.
XX
AC ADA08844;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide #95.
XX
KW ds; human; gene therapy; ovarian cancer; cancer.
XX
OS Homo sapiens.
XX
PN US2003091580-A1.
XX
PD 15-MAY-2003.
XX
PF 17-JUL-2001; 2001US-00907969.
XX
PR 18-JUN-2001; 2001US-00884441.
XX
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGA/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (HILL/) HILL P.
PA (ALBO/) ALBONE E.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2003-532352/50.
XX
PT New isolated O772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT cancer.
XX
PS Example 1; Fig 15; 371pp; English.
XX
CC The invention relates to an isolated O772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma cDNAs and protein cDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma
CC antigen polynucleotide.
XX
SQ Sequence 605 BP; 111 A; 144 C; 140 G; 159 T; 0 U; 51 Other;

Query Match 8.7%; Score 191; DB 9; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
                ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 117
ADF08591/c
ID ADF08591 standard; cDNA; 605 BP.
XX
AC ADF08591;
XX
DT 12-FEB-2004 (first entry)
XX
DE cDNA encoding secreted ovarian carcinoma antigen seqid 95.
XX
KW gene therapy; protein therapy; vaccine; antibody inhibition;
KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW secreted ovarian carcinoma antigen; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003124140-A1.
XX
PD 03-JUL-2003.
XX
PF 17-JUL-2002; 2002US-00198053.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2003-897152/82.
XX
PT Oncogenic nucleic acids useful for the prevention, diagnosis and
PT treatment of breast cancer.
XX
PS Example 1; SEQ ID NO 95; 399pp; English.
XX
CC The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of (II) by expressing
CC inactive proteins or to supplement the patients own production of (II).
CC Additionally, (I) may be used to produce (II), by inserting (I) into a
CC host cell and culturing the cell to express the protein (II). (I) And its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative

```
CC   therapy. The host cell may also be used as antigens in the production of
CC   antibodies against (II) and in assays to identify modulators of (II)'s
CC   expression and activity. The anti-(II) antibodies, agonists and
CC   antagonists may be used to regulate expression and activity and as
CC   diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC   immunoassay). This sequence represents a polynucleotide encoding a
CC   secreted ovarian carcinoma antigen.
XX
SQ   Sequence 605 BP; 111 A; 144 C; 140 G; 159 T; 0 U; 51 Other;

  Query Match             8.7%;  Score 191;  DB 10;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 118
ADG46339/c
ID   ADG46339 standard; cDNA; 605 BP.
XX
AC   ADG46339;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human ovarian carcinoma polynucleotide #95.
XX
KW   Human; ovarian carcinoma; gene; ss; O8E; ovarian cancer;
KW   secreted tumour antigen; cytostatic; O772P.
XX
OS   Homo sapiens.
XX
PN   US2003165504-A1.
XX
PD   04-SEP-2003.
XX
PF   04-APR-2001; 2001US-00827271.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
XX
PA   (RETT/) RETTER M W.
PA   (FANG/) FANGER G R.
XX
PI   Retter MW,  Fanger GR;
XX
DR   WPI; 2003-898035/82.
XX
PT   New isolated O8E or O772P polypeptides, useful for diagnosing,
PT   preventing, treating and monitoring cancer, e.g. ovarian cancer,
PT   stimulating the immune response in patient.
XX
PS   Example 1; SEQ ID NO 95; 290pp; English.
XX
CC   The invention relates to human ovarian carcinoma polypeptides, designated
CC   O8E or O772P, and the polynucleotides encoding them. The invention also
CC   relates to methods for inhibiting the development of cancer, e.g. ovarian
CC   cancer in a patient, methods for stimulating and/or expanding T cells and
CC   methods for identifying secreted tumour antigens. The polypeptides,
CC   compositions, antibodies to the polypeptides and methods are useful for
CC   diagnosing, preventing, treating and monitoring cancer, e.g. ovarian
CC   cancer. The composition is particularly useful for stimulating an immune
CC   response in patient. This sequence represents a human ovarian carcinoma
CC   polynucleotide of the invention.
XX
SQ   Sequence 605 BP; 111 A; 144 C; 140 G; 159 T; 0 U; 51 Other;

  Query Match             8.7%;  Score 191;  DB 10;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 119
ADX17474/c
ID   ADX17474 standard; cDNA; 605 BP.
XX
AC   ADX17474;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human ovarian carcinoma antigen cDNA #88.
XX
KW   diagnosis; cytostatic; immunostimulant; gene therapy; tumor;
KW   ovarian tumor; cancer; carcinoma; antigen; ss.
XX
OS   Homo sapiens.
XX
PN   US2005031634-A1.
XX
PD   10-FEB-2005.
XX
PF   02-JUN-2004; 2004US-00860790.
XX
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
PR   04-APR-2001; 2001US-00827271.
PR   18-JUN-2001; 2001US-00884441.
PR   17-JUL-2001; 2001US-00907969.
PR   17-JUL-2002; 2002US-00198053.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Bangur CS,  Retter MW,  Fanger GR,  Hill P;
XX
DR   WPI; 2005-151645/16.
XX
PT   New ovarian carcinoma polynucleotides, preferably cDNAs, useful for
```

PT diagnosing, preventing and treating diseases, such as ovarian cancer, and
PT for eliciting humoral and/or cellular immune response.
XX
PS Example 1; SEQ ID NO 95; 398pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising a sequence
CC of, a sequence hybridizing under highly stringent conditions to, or
CC having at least 75, or 90% identity to a 849 or 1399 bp, fully defined in
CC the specification, its complement or degenerate variants, or a sequence
CC of at least 20 contiguous residues of the 849 or 1399 bp sequence. The
CC polynucleotides and polypeptides are useful for diagnosing, preventing
CC and treating diseases, such as ovarian cancer, and for eliciting humoral
CC and/or cellular immune response. This sequence corresponds to an ovarian
CC carcinoma antigen cDNA of the invention.
XX
SQ Sequence 605 BP; 111 A; 144 C; 140 G; 159 T; 0 U; 51 Other;

Query Match 8.7%; Score 191; DB 14; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
Db 359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy 1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
Db 299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy 2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
Db 239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy 2079 AAAGAGTCCAC 2089
Db 179 AAAGAGTCCAC 169

RESULT 120
AAA69790
ID AAA69790 standard; cDNA; 697 BP.
XX
AC AAA69790;
XX
DT 07-NOV-2000 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide SEQ ID NO:100.
XX
KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200036107-A2.
XX
PD 22-JUN-2000.
XX
PF 17-DEC-1999; 99WO-US030270.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TN;
XX
DR WPI; 2000-431589/37.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT encoding it, useful for the diagnosis, prevention and treatment of
PT cancer, preferably ovarian cancer.
XX
PS Claim 18; Fig 15; 299pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557 represent human
CC ovarian carcinoma polynucleotides and proteins used in the
CC exemplification of the present invention
XX
SQ Sequence 697 BP; 183 A; 169 C; 169 G; 130 T; 0 U; 46 Other;

Query Match 8.7%; Score 191; DB 3; Length 697;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
Db 342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy 1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
Db 402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy 2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
Db 462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy 2079 AAAGAGTCCAC 2089
Db 522 AAAGAGTCCAC 532

RESULT 121
ABN72684
ID ABN72684 standard; DNA; 697 BP.
XX
AC ABN72684;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ovarian carcinoma antigen polynucleotide #100.
XX
KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US022635.
XX
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2002-164781/21.
XX
PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX

PS Example 1; Page 133; 408pp; English.
XX
CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents DNA related
CC to the invention
XX
SQ Sequence 697 BP; 183 A; 169 C; 169 G; 130 T; 0 U; 46 Other;

Query Match 8.7%; Score 191; DB 6; Length 697;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532
```

RESULT 122
ADA08849
ID ADA08849 standard; DNA; 697 BP.
XX
AC ADA08849;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human ovarian carcinoma antigen polynucleotide #100.
XX
KW ds; human; gene therapy; ovarian cancer; cancer.
XX
OS Homo sapiens.
XX
PN US2003091580-A1.
XX
PD 15-MAY-2003.
XX
PF 17-JUL-2001; 2001US-00907969.
XX
PR 18-JUN-2001; 2001US-00884441.
XX
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
PA (ALGA/) ALGATE P A.
PA (FLIN/) FLING S P.
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
PA (REED/) REED S G.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (HILL/) HILL P.
PA (ALBO/) ALBONE E.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
DR WPI; 2003-532352/50.
XX
PT New isolated O772P polypeptides and polynucleotides, useful in gene
PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT cancer.
XX
PS Example 1; Fig 15; 371pp; English.
XX
CC The invention relates to an isolated O772P polypeptide, which has the
CC structure fully defined in the specification. The composition containing
CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC or antigen presenting cells are useful for stimulating an immune response
CC and treating ovarian cancer. Detecting the presence of the
CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC carcinoma cDNAs and protein cDNAs were identified using microarray
CC technology. The present sequence represents a human ovarian carcinoma
CC antigen polynucleotide.
XX
SQ Sequence 697 BP; 183 A; 169 C; 169 G; 130 T; 0 U; 46 Other;

Query Match 8.7%; Score 191; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532
```

RESULT 123
ADF08596
ID ADF08596 standard; cDNA; 697 BP.
XX
AC ADF08596;
XX
DT 12-FEB-2004 (first entry)
XX
DE cDNA encoding secreted ovarian carcinoma antigen seqid 100.
XX
KW gene therapy; protein therapy; vaccine; antibody inhibition;
KW breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW secreted ovarian carcinoma antigen; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003124140-A1.
XX
PD 03-JUL-2003.
XX
PF 17-JUL-2002; 2002US-00198053.
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
XX
PA (CORI-) CORIXA CORP.
XX

```
PI   Bangur CS,  Retter MW,  Fanger GR,  Hill P;
XX
DR   WPI; 2003-897152/82.
XX
PT   Oncogenic nucleic acids useful for the prevention, diagnosis and
PT   treatment of breast cancer.
XX
PS   Example 1; SEQ ID NO 100; 399pp; English.
XX
CC   The invention describes nucleic acids (I) and the polypeptides (II) they
CC   encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC   treating diseases related to their aberrant expression i.e. breast
CC   cancers. For example, (I) and (II) may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of (II) by expressing
CC   inactive proteins or to supplement the patients own production of (II).
CC   Additionally, (I) may be used to produce (II), by inserting (I) into a
CC   host cell and culturing the cell to express the protein (II). (I) And its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids in
CC   samples, and therefore which patients may be in need of restorative
CC   therapy. The host cell may also be used as antigens in the production of
CC   antibodies against (II) and in assays to identify modulators of (II)'s
CC   expression and activity. The anti-(II) antibodies, agonists and
CC   antagonists may be used to regulate expression and activity and as
CC   diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC   immunoassay). This sequence represents a polynucleotide encoding a
CC   secreted ovarian carcinoma antigen.
XX
SQ   Sequence 697 BP; 183 A; 169 C; 169 G; 130 T; 0 U; 46 Other;

  Query Match             8.7%;  Score 191;  DB 10;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 124
ADG46344
ID   ADG46344 standard; cDNA; 697 BP.
XX
AC   ADG46344;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human ovarian carcinoma polynucleotide #100.
XX
KW   Human; ovarian carcinoma; gene; ss; O8E; ovarian cancer;
KW   secreted tumour antigen; cytostatic; O772P.
XX
OS   Homo sapiens.
XX
PN   US2003165504-A1.
XX
PD   04-SEP-2003.
XX
PF   04-APR-2001; 2001US-00827271.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
XX
PA   (RETT/) RETTER M W.
PA   (FANG/) FANGER G R.
XX
PI   Retter MW,  Fanger GR;
XX
DR   WPI; 2003-898035/82.
XX
PT   New isolated O8E or O772P polypeptides, useful for diagnosing,
PT   preventing, treating and monitoring cancer, e.g. ovarian cancer,
PT   stimulating the immune response in patient.
XX
PS   Example 1; SEQ ID NO 100; 290pp; English.
XX
CC   The invention relates to human ovarian carcinoma polypeptides, designated
CC   O8E or O772P, and the polynucleotides encoding them. The invention also
CC   relates to methods for inhibiting the development of cancer, e.g. ovarian
CC   cancer in a patient, methods for stimulating and/or expanding T cells and
CC   methods for identifying secreted tumour antigens. The polypeptides,
CC   compositions, antibodies to the polypeptides and methods are useful for
CC   diagnosing, preventing, treating and monitoring cancer, e.g. ovarian
CC   cancer. The composition is particularly useful for stimulating an immune
CC   response in patient. This sequence represents a human ovarian carcinoma
CC   polynucleotide of the invention.
XX
SQ   Sequence 697 BP; 183 A; 169 C; 169 G; 130 T; 0 U; 46 Other;

  Query Match             8.7%;  Score 191;  DB 10;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 125
ADX17479
ID   ADX17479 standard; cDNA; 697 BP.
XX
AC   ADX17479;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human ovarian carcinoma antigen cDNA #93.
XX
KW   diagnosis; cytostatic; immunostimulant; gene therapy; tumor;
KW   ovarian tumor; cancer; carcinoma; antigen; ss.
XX
OS   Homo sapiens.
XX
PN   US2005031634-A1.
XX
PD   10-FEB-2005.
```

XX
PF 02-JUN-2004; 2004US-00860790.
XX
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
PR 17-JUL-2002; 2002US-00198053.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2005-151645/16.
XX
PT New ovarian carcinoma polynucleotides, preferably cDNAs, useful for
PT diagnosing, preventing and treating diseases, such as ovarian cancer, and
PT for eliciting humoral and/or cellular immune response.
XX
PS Example 1; SEQ ID NO 100; 398pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising a sequence
CC of, a sequence hybridizing under highly stringent conditions to, or
CC having at least 75, or 90% identity to a 849 or 1399 bp, fully defined in
CC the specification, its complement or degenerate variants, or a sequence
CC of at least 20 contiguous residues of the 849 or 1399 bp sequence. The
CC polynucleotides and polypeptides are useful for diagnosing, preventing
CC and treating diseases, such as ovarian cancer, and for eliciting humoral
CC and/or cellular immune response. This sequence corresponds to an ovarian
CC carcinoma antigen cDNA of the invention.
XX
SQ Sequence 697 BP; 183 A; 169 C; 169 G; 130 T; 0 U; 46 Other;

Query Match 8.7%; Score 191; DB 14; Length 697;
Best Local Similarity 100.0%; Pred. No. 2.7e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
Db 342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy 1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
Db 402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy 2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
Db 462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy 2079 AAAGAGTCCAC 2089
Db 522 AAAGAGTCCAC 532

RESULT 126
AAD13071
ID AAD13071 standard; DNA; 6340 BP.
XX
AC AAD13071;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-1XUb-GFP construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; ubiquitin;
KW green fluorescent protein; GFP; human; cyclic; circular;
KW plasmid pcDNA3-1XUb-GFP; fusion protein; ds.
XX
OS Aequorea victoria.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 3; Page 125-126; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-1XUb-GFP construct comprising mutant multimerised
CC human ubiquitin destabilisation domain-Aequorea victoria green
CC fluorescent protein (GFP) fusion protein. (Updated on 11-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 6340 BP; 1482 A; 1706 C; 1634 G; 1518 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 6340;
Best Local Similarity 100.0%; Pred. No. 8.2e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
Db 925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
Db 985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
Db 1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy 1638 CACCTGGT 1645
Db 1105 CACCTGGT 1112

RESULT 127
AAD13063
ID AAD13063 standard; DNA; 6411 BP.
XX

```
AC   AAD13063;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-Bla;
KW   fusion protein; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 2; Page 116-117; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-Bla construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Escherichia coli beta-lactamase
CC   fusion protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6411 BP; 1522 A; 1644 C; 1647 G; 1598 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6411;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
              ||||||||
Db       1105 CACCTGGT 1112


RESULT 128
AAD13067
ID   AAD13067 standard; DNA; 6411 BP.
XX
AC   AAD13067;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-Ub-Met-Bla;
KW   fusion protein; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 2; Page 122-123; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-Ub-Met-Bla construct comprising wild type
CC   multimerised human ubiquitin destabilisation domain-Escherichia coli beta
CC   -lactamase fusion protein. (Updated on 11-SEP-2003 to standardise OS
CC   field)
XX
SQ   Sequence 6411 BP; 1522 A; 1642 C; 1649 G; 1598 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6411;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 129
AAD13078
ID   AAD13078 standard; DNA; 6436 BP.
XX
AC   AAD13078;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-C3;
KW   fusion protein; ds.
XX
OS   Mammalia.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 4; Page 133-134; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
CC   (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6436 BP; 1577 A; 1606 C; 1607 G; 1646 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6436;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 130
AAD13087
ID   AAD13087 standard; DNA; 6459 BP.
XX
AC   AAD13087;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-DEVD-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-DEVD-Bla;
KW   fusion protein; caspase-3; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 10; Page 142-143; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
```

```
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-DEVD-Bla construct comprising mutant
CC   multimerised human ubiquitin destabilisation domain-Escherichia coli beta
CC   -lactamase fusion protein linked by a linker containing a caspase-3
CC   cleavage site. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6459 BP; 1527 A; 1659 C; 1668 G; 1605 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6459;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 131
AAD13093
ID   AAD13093 standard; DNA; 6459 BP.
XX
AC   AAD13093;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-1XUb-DEVA-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-1XUb-DEVA-Bla;
KW   fusion protein; caspase-3; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 10; Page 149-150; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-1XUb-DEVA-Bla construct comprising mutant
CC   multimerised human ubiquitin destabilisation domain-Escherichia coli beta
CC   -lactamase fusion protein linked by a linker containing a DEVA site that
CC   should not serve as a cleavage site for caspase-3-like proteases.
CC   (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6459 BP; 1526 A; 1659 C; 1668 G; 1606 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6459;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112


RESULT 132
AAD13072
ID   AAD13072 standard; DNA; 6607 BP.
XX
AC   AAD13072;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-GFP construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; ubiquitin;
KW   green fluorescent protein; GFP; human; cyclic; circular;
KW   plasmid pcDNA3-2XUb-GFP; fusion protein; ds.
XX
OS   Aequorea victoria.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
```

XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 3; Page 126-127; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-2XUb-GFP construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Aequorea victoria green
CC fluorescent protein (GFP) fusion protein. (Updated on 11-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 6607 BP; 1558 A; 1777 C; 1701 G; 1571 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 6607;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136
```

RESULT 133
AAD13064
ID AAD13064 standard; DNA; 6678 BP.
XX
AC AAD13064;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-2XUb-Bla construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-Bla;
KW fusion protein; ds.
XX
OS Escherichia coli.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 2; Page 117-118; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-2XUb-Bla construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Escherichia coli beta-lactamase
CC fusion protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 6678 BP; 1598 A; 1715 C; 1714 G; 1651 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 6678;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136
```

RESULT 134
AAD13079
ID AAD13079 standard; DNA; 6703 BP.
XX
AC AAD13079;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX

```
DE   pcDNA3-2XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-C3;
KW   fusion protein; ds.
XX
OS   Mammalia.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 4; Page 134-135; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3 fusion
CC   protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6703 BP; 1653 A; 1677 C; 1674 G; 1699 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6703;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136
```

```
RESULT 135
AAD13088
ID   AAD13088 standard; DNA; 6726 BP.
XX
AC   AAD13088;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-DEVD-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-DEVD-Bla;
KW   fusion protein; caspase-3; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 10; Page 144-145; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-DEVD-Bla construct comprising mutant
CC   multimerised human ubiquitin destabilisation domains-Escherichia coli
CC   beta-lactamase fusion protein linked by a linker containing a caspase-3
CC   cleavage site. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6726 BP; 1603 A; 1730 C; 1735 G; 1658 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6726;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068
```

```
Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 136
AAD13094
ID   AAD13094 standard; DNA; 6727 BP.
XX
AC   AAD13094;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-2XUb-DEVA-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-2XUb-DEVA-Bla;
KW   fusion protein; caspase-3; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 10; Page 151-152; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-2XUb-DEVA-Bla construct comprising mutant
CC   multimerised human ubiquitin destabilisation domain-Escherichia coli beta
CC   -lactamase fusion protein linked by a linker containing a DEVA site that
CC   should not serve as a cleavage site for caspase-3-like proteases.
CC   (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6727 BP; 1602 A; 1731 C; 1735 G; 1659 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6727;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1009

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1069

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1129

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1130 CACCTGGT 1137


RESULT 137
AAD13073
ID   AAD13073 standard; DNA; 6850 BP.
XX
AC   AAD13073;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-GFP construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; ubiquitin;
KW   green fluorescent protein; GFP; human; cyclic; circular;
KW   plasmid pcDNA3-3XUb-GFP; fusion protein; ds.
XX
OS   Aequorea victoria.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 3; Page 128-129; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
```

```
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-GFP construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Aequorea victoria green
CC   fluorescent protein (GFP) fusion protein. (Updated on 11-SEP-2003 to
CC   standardise OS field)
XX
SQ   Sequence 6850 BP; 1628 A; 1842 C; 1762 G; 1618 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6850;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 138
AAD13065
ID   AAD13065 standard; DNA; 6921 BP.
XX
AC   AAD13065;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla;
KW   fusion protein; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 2; Page 119-120; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-Bla construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Escherichia coli beta-lactamase
CC   fusion protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6921 BP; 1668 A; 1780 C; 1775 G; 1698 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6921;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 139
AAD13080
ID   AAD13080 standard; DNA; 6946 BP.
XX
AC   AAD13080;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-C3;
KW   fusion protein; ds.
XX
OS   Mammalia.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
```

```
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 4; Page 136-137; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-C3 construct comprising mutant multimerised
CC   human ubiquitin destabilisation domains-Mammalia caspase-3 fusion
CC   protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6946 BP; 1723 A; 1742 C; 1735 G; 1746 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6946;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 140
AAD13095
ID   AAD13095 standard; DNA; 6967 BP.
XX
AC   AAD13095;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-DEVA-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-DEVA-Bla;
KW   fusion protein; caspase-3; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 10; Page 152-153; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-DEVA-Bla construct comprising mutant
CC   multimerised human ubiquitin destabilisation domain-Escherichia coli beta
CC   -lactamase fusion protein linked by a linker containing a DEVA site that
CC   should not serve as a cleavage site for caspase-3-like proteases.
CC   (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 6967 BP; 1672 A; 1795 C; 1796 G; 1704 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 6967;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 141
AAD13089
ID   AAD13089 standard; DNA; 6969 BP.
XX
AC   AAD13089;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-DEVD-Bla construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
```

KW transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-DEVD-Bla;
KW fusion protein; caspase-3; ds.
XX
OS Escherichia coli.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 10; Page 145-146; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-3XUb-DEVD-Bla construct comprising mutant
CC multimerised human ubiquitin destabilisation domains-Escherichia coli
CC beta-lactamase fusion protein linked by a linker containing a caspase-3
CC cleavage site. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 6969 BP; 1673 A; 1795 C; 1796 G; 1705 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 6969;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy 1638 CACCTGGT 1645
||||||||
Db 1129 CACCTGGT 1136

RESULT 142
AAD13110
ID AAD13110 standard; DNA; 7051 BP.
XX
AC AAD13110;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-Ub-Met-Bla HRV16(C106A) construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW human; ubiquitin; cyclic; circular; rhinovirus 16 2A protease;
KW plasmid pcDNA3-Ub-Met-Bla HRV16(C106A); fusion protein; ds.
XX
OS Escherichia coli.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 13; Page 167-168; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-Ub-Met-Bla HRV16(C106A) construct comprising
CC multimerised human wild type ubiquitin destabilisation domains-
CC Escherichia coli beta-lactamase-human mutant (C106A) rhinovirus 16 2A
CC protease fusion protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 7051 BP; 1746 A; 1765 C; 1770 G; 1770 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7051;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 926 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 985

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 986 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1045

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db       1046 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1105

Qy       1638 CACCTGGT 1645
              ||||||||
Db       1106 CACCTGGT 1113


RESULT 143
AAD13111
ID   AAD13111 standard; DNA; 7052 BP.
XX
AC   AAD13111;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-Bla HRV16(D35A) construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; rhinovirus 16 2A protease;
KW   plasmid pcDNA3-Ub-Met-Bla HRV16(D35A); fusion protein; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 13; Page 169-170; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-Ub-Met-Bla HRV16(D35A) construct comprising
CC   multimerised human wild type ubiquitin destabilisation domains-
CC   Escherichia coli beta-lactamase-human mutant (D35A) rhinovirus 16 2A
CC   protease fusion protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 7052 BP; 1744 A; 1767 C; 1770 G; 1771 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 7052;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
              ||||||||
Db       1105 CACCTGGT 1112


RESULT 144
AAD13105
ID   AAD13105 standard; DNA; 7053 BP.
XX
AC   AAD13105;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-Bla HRV16 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla HRV14;
KW   fusion protein; rhinovirus 16 2A protease; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 13; Page 162-163; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-Bla HRV16 construct comprising wildtype
```

CC multimerised human ubiquitin destabilisation domains-Escherichia coli
CC beta-lactamase-human rhinovirus 16 2A protease fusion protein. (Updated
CC on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 7053 BP; 1746 A; 1766 C; 1770 G; 1771 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7053;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy 1638 CACCTGGT 1645
||||||||
Db 1105 CACCTGGT 1112

RESULT 145
AAD13074
ID AAD13074 standard; DNA; 7093 BP.
XX
AC AAD13074;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-4XUb-GFP construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; ubiquitin;
KW green fluorescent protein; GFP; human; cyclic; circular;
KW plasmid pcDNA3-4XUb-GFP; fusion protein; ds.
XX
OS Aequorea victoria.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 3; Page 130-131; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-4XUb-GFP construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Aequorea victoria green
CC fluorescent protein (GFP) fusion protein. (Updated on 11-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 7093 BP; 1698 A; 1907 C; 1823 G; 1665 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7093;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy 1638 CACCTGGT 1645
||||||||
Db 1129 CACCTGGT 1136

RESULT 146
AAD13066
ID AAD13066 standard; DNA; 7164 BP.
XX
AC AAD13066;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-4XUb-Bla construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-Bla;
KW fusion protein; ds.
XX
OS Escherichia coli.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to

PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 2; Page 120-121; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-4XUb-Bla construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Escherichia coli beta-lactamase
CC fusion protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 7164 BP; 1738 A; 1845 C; 1836 G; 1745 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7164;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy 1638 CACCTGGT 1645
||||||||
Db 1129 CACCTGGT 1136

RESULT 147
AAD13081
ID AAD13081 standard; DNA; 7189 BP.
XX
AC AAD13081;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-4XUb-C3 construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; caspase-3;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-C3;
KW fusion protein; ds.
XX
OS Mammalia.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 4; Page 137-138; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-4XUb-C3 construct comprising mutant multimerised
CC human ubiquitin destabilisation domains-Mammalia caspase-3 fusion
CC protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 7189 BP; 1793 A; 1807 C; 1796 G; 1793 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7189;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy 1638 CACCTGGT 1645
||||||||
Db 1129 CACCTGGT 1136

RESULT 148
AAD13096
ID AAD13096 standard; DNA; 7212 BP.
XX
AC AAD13096;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-4XUb-DEVA-Bla construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-DEVA-Bla;
KW fusion protein; caspase-3; ds.
XX
OS Escherichia coli.

OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 10; Page 154-155; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-4XUb-DEVA-Bla construct comprising mutant
CC multimerised human ubiquitin destabilisation domain-Escherichia coli beta
CC -lactamase fusion protein linked by a linker containing a DEVA site that
CC should not serve as a cleavage site for caspase-3-like proteases.
CC (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 7212 BP; 1742 A; 1860 C; 1857 G; 1753 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7212;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
Db 949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
Db 1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
Db 1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy 1638 CACCTGGT 1645
Db 1129 CACCTGGT 1136

RESULT 149
AAD13090
ID AAD13090 standard; DNA; 7212 BP.
XX
AC AAD13090;
XX
DT 11-SEP-2003 (revised)
DT 16-OCT-2001 (first entry)
XX
DE pcDNA3-4XUb-DEVD-Bla construct.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW human; ubiquitin; cyclic; circular; plasmid pcDNA3-4XUb-DEVD-Bla;
KW fusion protein; caspase-3; ds.
XX
OS Escherichia coli.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 10; Page 147-148; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-4XUb-DEVD-Bla construct comprising mutant
CC multimerised human ubiquitin destabilisation domains-Escherichia coli
CC beta-lactamase fusion protein linked by a linker containing a caspase-3
CC cleavage site. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 7212 BP; 1743 A; 1859 C; 1857 G; 1753 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 7212;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
Db 948 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1007

Qy 1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
Db 1008 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1067

Qy 1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
Db 1068 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1127

Qy 1638 CACCTGGT 1645

```
Db       1128 CACCTGGT 1135


RESULT 150
AAD13084
ID   AAD13084 standard; DNA; 7247 BP.
XX
AC   AAD13084;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-C3 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; caspase-3;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-Ub-Met-C3;
KW   fusion protein; ds.
XX
OS   Mammalia.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 4; Page 140-141; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-Ub-Met-C3 construct comprising wildtype multimerised
CC   human ubiquitin destabilisation domain-Mammalia caspase-3 fusion protein.
CC   (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 7247 BP; 1794 A; 1796 C; 1826 G; 1831 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 7247;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
              ||||||||
Db       1105 CACCTGGT 1112


RESULT 151
AAD13112
ID   AAD13112 standard; DNA; 7511 BP.
XX
AC   AAD13112;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-Ub-Met-Bla HRV14 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla HRV14;
KW   fusion protein; rhinovirus 14 2A protease; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 13; Page 170-171; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-Bla HRV14 construct comprising wildtype
CC   multimerised human ubiquitin destabilisation domains-Escherichia coli
CC   beta-lactamase-human rhinovirus 14 2A protease fusion protein. (Updated
CC   on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 7511 BP; 1884 A; 1853 C; 1893 G; 1881 T; 0 U; 0 Other;
```

```
  Query Match             8.6%;  Score 188;  DB 4;  Length 7511;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 984

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         985 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1044

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1104

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1105 CACCTGGT 1112
```

```
RESULT 152
AAD13108
ID   AAD13108 standard; DNA; 7563 BP.
XX
AC   AAD13108;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-Bla HRV16(C106A) construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla HRV16(C106A);
KW   fusion protein; rhinovirus 16 2A protease; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 13; Page 164-165; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-Bla HRV16(C106A) construct comprising
CC   multimerised human mutant ubiquitin destabilisation domains- Escherichia
CC   coli beta-lactamase-human mutant (C106A) rhinovirus 16 2A protease fusion
CC   protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 7563 BP; 1892 A; 1905 C; 1896 G; 1870 T; 0 U; 0 Other;
```

```
  Query Match             8.6%;  Score 188;  DB 4;  Length 7563;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136
```

```
RESULT 153
AAD13104
ID   AAD13104 standard; DNA; 7563 BP.
XX
AC   AAD13104;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-Bla HRV16 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla HRV16;
KW   fusion protein; rhinovirus 16 2A protease; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 13; Page 160-161; 171pp; English.
XX
```

```
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-Bla HRV16 construct comprising mutant
CC   multimerised human ubiquitin destabilisation domains-Escherichia coli
CC   beta-lactamase-human rhinovirus 16 2A protease fusion protein. (Updated
CC   on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 7563 BP; 1892 A; 1904 C; 1896 G; 1871 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 7563;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 154
AAD13109
ID   AAD13109 standard; DNA; 7564 BP.
XX
AC   AAD13109;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-Bla HRV16(D35A) construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla HRV16(D35A);
KW   fusion protein; rhinovirus 16 2A protease; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO200157242-A2.
XX
PD   09-AUG-2001.
XX
PF   02-FEB-2001; 2001WO-US003791.
XX
PR   04-FEB-2000; 2000US-00498098.
XX
PA   (AURO-) AURORA BIOSCIENCES CORP.
XX
PI   Stack JH,  Whitney M,  Cubitt AB,  Pollok BA;
XX
DR   WPI; 2001-488890/53.
XX
PT   Destabilizing proteins in living cells, by coupling a target protein to
PT   linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT   protein endoproteases, comprising two copies of the domain.
XX
PS   Example 13; Page 166-167; 171pp; English.
XX
CC   The present invention relates to a method for destabilising a target
CC   protein in a cell. The method comprises a linker moiety which operatively
CC   couples a target protein (a reporter moiety) to a linear multimerised
CC   destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC   endoproteases. The method is useful for detecting an activity such as
CC   protease, protein kinase or phosphoprotein phosphatase activity and is
CC   also useful for identifying modulators of these activities. The method is
CC   also useful for developing novel assays for a wide range of post-
CC   translational activities, such as proteolysis, phosphorylation,
CC   dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC   disulfide bond formation and ADP-ribosylation within cells. The
CC   recombinant DNA molecule of the invention is useful for creating
CC   transgenic animals useful as disease models and transgenic plants with
CC   improved disease resistance or other favourable traits. The present
CC   sequence is a pcDNA3-3XUb-Bla HRV16(D35A) construct comprising
CC   multimerised human mutant ubiquitin destabilisation domains- Escherichia
CC   coli beta-lactamase-human mutant (D35A) rhinovirus 16 2A protease fusion
CC   protein. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 7564 BP; 1892 A; 1905 C; 1896 G; 1871 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 188;  DB 4;  Length 7564;
  Best Local Similarity   100.0%;  Pred. No. 8.1e-83;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136


RESULT 155
AAD13100
ID   AAD13100 standard; DNA; 8022 BP.
XX
AC   AAD13100;
XX
DT   11-SEP-2003  (revised)
DT   16-OCT-2001  (first entry)
XX
DE   pcDNA3-3XUb-Bla HRV14 construct.
XX
KW   Protein destabilisation; linker moiety; reporter moiety; disease model;
KW   linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW   transgenic animal; transgenic plant; disease resistance; beta-lactamase;
KW   human; ubiquitin; cyclic; circular; plasmid pcDNA3-3XUb-Bla HRV14;
KW   fusion protein; rhinovirus 14 2A protease; ds.
XX
OS   Escherichia coli.
OS   Homo sapiens.
OS   Chimeric.
XX
```

PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 13; Page 157-158; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a pcDNA3-3XUb-Bla HRV14 construct comprising mutant
CC multimerised human ubiquitin destabilisation domains-Escherichia coli
CC beta-lactamase-human rhinovirus 14 2A protease fusion protein. (Updated
CC on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 8022 BP; 2030 A; 1992 C; 2019 G; 1981 T; 0 U; 0 Other;

Query Match 8.6%; Score 188; DB 4; Length 8022;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1008

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1068

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1128

Qy       1638 CACCTGGT 1645
             ||||||||
Db       1129 CACCTGGT 1136
```

RESULT 156
ABZ19630
ID ABZ19630 standard; cDNA; 471 BP.
XX
AC ABZ19630;
XX
DT 23-JAN-2003 (first entry)
XX
DE Group III cDNA cancer related clone SEQ ID NO:2056.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US010421.
XX
PR 30-MAR-2001; 2001US-0280255P.
PR 28-AUG-2001; 2001US-0315563P.
PR 09-JAN-2002; 2002US-0347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX
PS Claim 1; SEQ ID NO 2056; 207pp; English.
XX
CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 471 BP; 107 A; 120 C; 142 G; 101 T; 0 U; 1 Other;

Query Match 8.2%; Score 179; DB 8; Length 471;
Best Local Similarity 100.0%; Pred. No. 2.5e-78;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 316

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 376

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCT 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCT 435
```

RESULT 157
AAS24867
ID AAS24867 standard; cDNA; 227 BP.
XX
AC AAS24867;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #1048.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX

OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003733.
XX
PR 04-FEB-2000; 2000US-0180403P.
PR 28-MAR-2000; 2000US-0192745P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2001-488879/53.
XX
PT New polynucleotides encoding ovarian tumor proteins, useful for treating
PT ovarian cancer, and as probes, primers, and markers of cancer
PT progression.
XX
PS Example 1; Page 276; 378pp; English.
XX
CC The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumour
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumour polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumour DNA or protein by incubating isolated T-
CC cells allowing them to proliferate, and administering to the patient. The
CC sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumour protein cDNA clones
XX
SQ Sequence 227 BP; 64 A; 60 C; 60 G; 43 T; 0 U; 0 Other;

Query Match 8.0%; Score 176; DB 4; Length 227;
Best Local Similarity 99.6%; Pred. No. 7.9e-77;
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 60

Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 TGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 120

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 180

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 227
```

RESULT 158
AAH83506
ID AAH83506 standard; cDNA; 335 BP.
XX
AC AAH83506;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:1130.
XX
KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200151513-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US001575.
XX
PR 14-JAN-2000; 2000US-0176722P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA;
XX
DR WPI; 2001-425866/45.
XX
PT Novel ovarian tumor proteins, and nucleic acids encoding them, used to
PT treat and diagnose cancers, particularly ovarian cancer.
XX
PS Claim 5; Page 270; 338pp; English.
XX
CC AAH82377 to AAH83878 represent human ovarian tumour-associated
CC polynucleotide sequences which encode ovarian tumour proteins. The
CC ovarian tumour protein and polynucleotide sequences have cytostatic
CC activity, and can be used in gene therapy and vaccine production. The
CC ovarian tumour proteins and polynucleotides can be used to inhibit the
CC development of cancer, particularly ovarian cancer. They can also be used
CC to diagnose the onset and progression of cancer
XX
SQ Sequence 335 BP; 84 A; 91 C; 92 G; 66 T; 0 U; 2 Other;

Query Match 8.0%; Score 176; DB 5; Length 335;
Best Local Similarity 99.6%; Pred. No. 7.9e-77;
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 60

Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 TGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 120

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 180

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 227
```

RESULT 159
ABZ19396
ID ABZ19396 standard; cDNA; 176 BP.
XX
AC ABZ19396;
XX
DT 23-JAN-2003 (first entry)
XX
DE Group III cDNA cancer related clone SEQ ID NO:1822.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX

PF 28-MAR-2002; 2002WO-US010421.
XX
PR 30-MAR-2001; 2001US-0280255P.
PR 28-AUG-2001; 2001US-0315563P.
PR 09-JAN-2002; 2002US-0347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX
PS Claim 1; SEQ ID NO 1822; 207pp; English.
XX
CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 176 BP; 47 A; 41 C; 52 G; 36 T; 0 U; 0 Other;

Query Match 7.8%; Score 171; DB 8; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.4e-74;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1185 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1244
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 65

Qy 1245 ATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 1304
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 66 ATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 125

Qy 1305 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1355
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 176

RESULT 160
ABV16078/c
ID ABV16078 standard; cDNA; 294 BP.
XX
AC ABV16078;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 16069.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 2691; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 294 BP; 67 A; 84 C; 72 G; 69 T; 0 U; 2 Other;

Query Match 7.8%; Score 171; DB 5; Length 294;
Best Local Similarity 100.0%; Pred. No. 2.4e-74;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 518 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 577
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 262 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 203

Qy 578 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCC 637
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 202 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCC 143

Qy 638 TGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142 TGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 92

RESULT 161
AAD13042
ID AAD13042 standard; cDNA; 228 BP.
XX
AC AAD13042;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human ubiquitin destabilising domain cDNA.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; human;
KW ubiquitin; ss.
XX
OS Homo sapiens.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.

XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Example 2; Page 95; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a human ubiquitin destabilising domain cDNA, used in the
CC exemplification of the invention
XX
SQ Sequence 228 BP; 65 A; 59 C; 61 G; 43 T; 0 U; 0 Other;

Query Match 7.6%; Score 167; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.3e-72;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1479 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 102

Qy       1539 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 162

Qy       1599 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 209
```

RESULT 162
AAD13057
ID AAD13057 standard; DNA; 228 BP.
XX
AC AAD13057;
XX
DT 16-OCT-2001 (first entry)
XX
DE DNA encoding human mutant ubiquitin (G76V) destabilising domain.
XX
KW Protein destabilisation; linker moiety; reporter moiety; disease model;
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;
KW transgenic animal; transgenic plant; disease resistance; human;
KW ubiquitin; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200157242-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US003791.
XX
PR 04-FEB-2000; 2000US-00498098.
XX
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Stack JH, Whitney M, Cubitt AB, Pollok BA;
XX
DR WPI; 2001-488890/53.
XX
PT Destabilizing proteins in living cells, by coupling a target protein to
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin
PT protein endoproteases, comprising two copies of the domain.
XX
PS Claim 8; Page 112; 171pp; English.
XX
CC The present invention relates to a method for destabilising a target
CC protein in a cell. The method comprises a linker moiety which operatively
CC couples a target protein (a reporter moiety) to a linear multimerised
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein
CC endoproteases. The method is useful for detecting an activity such as
CC protease, protein kinase or phosphoprotein phosphatase activity and is
CC also useful for identifying modulators of these activities. The method is
CC also useful for developing novel assays for a wide range of post-
CC translational activities, such as proteolysis, phosphorylation,
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,
CC disulfide bond formation and ADP-ribosylation within cells. The
CC recombinant DNA molecule of the invention is useful for creating
CC transgenic animals useful as disease models and transgenic plants with
CC improved disease resistance or other favourable traits. The present
CC sequence is a DNA encoding human mutant ubiquitin (G76V) destabilising
CC domain
XX
SQ Sequence 228 BP; 65 A; 59 C; 60 G; 44 T; 0 U; 0 Other;

Query Match 7.6%; Score 167; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 2.3e-72;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1479 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1538
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 102

Qy       1539 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1598
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 162

Qy       1599 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        163 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 209
```

RESULT 163
ADR23185
ID ADR23185 standard; DNA; 1581 BP.
XX
AC ADR23185;
XX
DT 04-NOV-2004 (first entry)
XX
DE Green fluorescent protein-beta-arrestin 2-ubiquitin 48 coding sequence.
XX
KW Arrestin; beta-arrestin 2; green fluorescent protein; ubiquitin;
KW G-protein coupled receptor; gene therapy; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .1581
FT /*tag= a
FT /product= "YGFP-beta arrestin 2-ubiquitin 48"
FT /partial
FT /note= "No start codon"
XX
PN WO2004067717-A2.

XX
PD 12-AUG-2004.
XX
PF 26-JAN-2004; 2004WO-US002029.
XX
PR 24-JAN-2003; 2003US-0442403P.
XX
PA (UYDU-) UNIV DUKE.
XX
PI Shenoy S, Lefkowitz RJ;
XX
DR WPI; 2004-580992/56.
DR P-PSDB; ADR23188.
XX
PT New modified arrestin comprising an arrestin and a ubiquitin moiety or
PT their biologically active fragment, useful for treating a human or a non-
PT human subject suffering from a GPCR-related disease.
XX
PS Claim 9; Fig 10; 90pp; English.
XX
CC The present sequence is the coding sequence for a chimeric polypeptide
CC comprising enhanced green fluorescent protein (EGFP), beta-arrestin 2 and
CC ubiquitin 48. The chimeric polypeptide is a claimed example of modified
CC arrestins of the invention that comprise an arrestin (or biologically
CC active fragment) and ubiquitin (or biologically active fragment) and an
CC optional label. The modified arrestin has an increased affinity for a G-
CC protein coupled receptor (GPCR) such that the arrestin remains associated
CC with the GPCR and traffics with the GPCR into endosomes and hence does
CC not dissociate at or near the plasma membrane. The modified arrestin
CC increases the internalisation of the GPCR. The invention provides the
CC modified arrestins, nucleic acids encoding them, expression vectors and
CC host (mammalian, bacterial, yeast, plant, insect or animal) cells. The
CC modified arrestins are used in a claimed method of screening compounds
CC and sample solutions for a GPCR agonist, antagonist, inverse agonist or
CC desensitization active compound. Claimed methods of treating a human or
CC non-human subject suffering from a GPCR-related disease comprise
CC administering a compound identified by this method or a nucleic acid
CC encoding a modified arrestin.
XX
SQ Sequence 1581 BP; 386 A; 467 C; 402 G; 326 T; 0 U; 0 Other;

Query Match 7.1%; Score 156; DB 13; Length 1581;
Best Local Similarity 99.5%; Pred. No. 6.6e-67;
Matches 206; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1372 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 1431

Qy 606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1432 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 1491

Qy 666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1492 AGACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 1551

Qy 726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
|||||||||||||||||||||||||||
Db 1552 CACCTGGTACTCCGTCTCAGAGGTGGG 1578

RESULT 164
ABQ59438/c
ID ABQ59438 standard; cDNA; 509 BP.
XX
AC ABQ59438;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:3133.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 509 BP; 117 A; 115 C; 114 G; 148 T; 0 U; 15 Other;

Query Match 6.9%; Score 152; DB 6; Length 509;
Best Local Similarity 100.0%; Pred. No. 6.5e-65;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 174 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 115

Qy 2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 114 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 55

Qy 2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
||||||||||||||||||||||||||||||||
Db 54 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 23

RESULT 165
ABZ19236/c
ID ABZ19236 standard; cDNA; 150 BP.
XX
AC ABZ19236;
XX
DT 23-JAN-2003 (first entry)
XX
DE Group III cDNA cancer related clone SEQ ID NO:1662.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;

KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US010421.
XX
PR 30-MAR-2001; 2001US-0280255P.
PR 28-AUG-2001; 2001US-0315563P.
PR 09-JAN-2002; 2002US-0347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
PT virology, immunology, microbiology, molecular biology and recombinant DNA
PT techniques.
XX
PS Claim 1; SEQ ID NO 1662; 207pp; English.
XX
CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
CC invention. (I) and (II) have cytostatic activity and can be used in gene
CC therapy and vaccines. (I), (II), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques. N.B. The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 150 BP; 28 A; 42 C; 35 G; 45 T; 0 U; 0 Other;

Query Match 6.8%; Score 150; DB 8; Length 150;
Best Local Similarity 100.0%; Pred. No. 6.6e-64;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1434 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 1493
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 150 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 91

Qy 1494 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1553
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 90 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 31

Qy 1554 CAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
|||||||||||||||||||||||||||||||
Db 30 CAGCAGAGGTTGATCTTTGCTGGGAAACAG 1

RESULT 166
AAT19287
ID AAT19287 standard; cDNA to mRNA; 176 BP.
XX
AC AAT19287;
XX
DT 28-JUN-1996 (first entry)
XX
DE Human gene signature HUMGS00304.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN WO9514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP001916.
XX
PR 12-NOV-1993; 93JP-00355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Single-stranded DNA for identifying gene signatures - isolated from 3'-
PT directed human cDNA library that reflects relative abundance of corresp.
PT mRNA in specific human tissues.
XX
PS Claim 1; Page 350; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp. double
CC -stranded DNA) which comprises one of the 7837 "GS" sequences given in
CC AAT19001-T26837 and which is able to hybridise to part of human genomic
CC DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were
CC obtained from 3'-directed cDNA libraries prepared from various human
CC tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using
CC poly(T) as the sole primer. Since the 3'- untranslated sequence is unique
CC to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise
CC with specific mRNAs. Each library is constructed so as to reflect
CC accurately the relative abundance of different mRNAs in the particular
CC tissue from which it was derived. The appearance frequency of a given GS
CC in a cDNA library can be determined (esp. using primers and probes
CC derived from the GS sequences) as a means of diagnosing abnormal cell
CC function or for recognising different cell types
XX
SQ Sequence 176 BP; 43 A; 42 C; 38 G; 53 T; 0 U; 0 Other;

Query Match 6.8%; Score 150; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 6.6e-64;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 83

Qy 2103 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 84 CTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 143

Qy 2163 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
|||||||||||||||||||||||||||||
Db 144 CTTTCCTTTCAATAAAGTTGTTGCATTCCC 173

RESULT 167
ADR12556
ID ADR12556 standard; DNA; 228 BP.
XX
AC ADR12556;
XX
DT 04-NOV-2004 (first entry)
XX
DE Gene vaccine nucleic acid #12.
XX
KW ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW antibacterial; haemostatic; protozoacide; antiinflammatory;
KW neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW intracellular parasite; protozoan parasite infection; malaria;
KW tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW Ebola hemorrhagic fever; Trypanosoma; Chagas disease;

KW Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW Herpes virus; severe acute respiratory syndrome.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .228
FT /*tag= a
XX
PN WO2004067040-A1.
XX
PD 12-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-JP000975.
XX
PR 31-JAN-2003; 2003JP-00023507.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12557.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein
PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 23; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 228 BP; 65 A; 63 C; 59 G; 41 T; 0 U; 0 Other;

Query Match 6.7%; Score 147; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 2e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 834 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 893
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 82 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 141

Qy 894 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 953
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 201

Qy 954 CACCTGGTGCTCCGTCTCAGAGGTGGG 980
|||||||||||||||||||||||||||
Db 202 CACCTGGTGCTCCGTCTCAGAGGTGGG 228

RESULT 168
ADS87009
ID ADS87009 standard; DNA; 228 BP.
XX
AC ADS87009;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 25.
XX
KW vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
DR P-PSDB; ADS87010.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Disclosure; SEQ ID NO 25; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC invention.
XX
SQ Sequence 228 BP; 65 A; 63 C; 59 G; 41 T; 0 U; 0 Other;

Query Match 6.7%; Score 147; DB 13; Length 228;
Best Local Similarity 100.0%; Pred. No. 2e-62;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 834 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 893
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 82 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 141

Qy 894 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 953
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 201

Qy 954 CACCTGGTGCTCCGTCTCAGAGGTGGG 980
|||||||||||||||||||||||||||
Db 202 CACCTGGTGCTCCGTCTCAGAGGTGGG 228

RESULT 169
ABQ55128
ID ABQ55128 standard; cDNA; 794 BP.
XX
AC ABQ55128;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HAOSJ27 cDNA, SEQ ID NO:1008.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
PD 03-JAN-2002.
XX
PF 07-JUN-2001; 2001WO-US018569.
XX
PR 07-JUN-2000; 2000US-0209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-147878/19.
DR P-PSDB; ABP42051.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.
XX
PS Claim 1; SEQ ID NO 1008; 2922pp; English.
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 794 BP; 213 A; 183 C; 209 G; 137 T; 0 U; 52 Other;

Query Match 6.4%; Score 141; DB 6; Length 794;
Best Local Similarity 100.0%; Pred. No. 1.9e-59;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 60

Qy 1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 120

Qy 1562 GTTGATCTTTGCTGGGAAACA 1582
|||||||||||||||||||||
Db 121 GTTGATCTTTGCTGGGAAACA 141

RESULT 170
AAS25026/c
ID AAS25026 standard; cDNA; 240 BP.
XX
AC AAS25026;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #1207.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003733.
XX
PR 04-FEB-2000; 2000US-0180403P.
PR 28-MAR-2000; 2000US-0192745P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2001-488879/53.
XX
PT New polynucleotides encoding ovarian tumor proteins, useful for treating
PT ovarian cancer, and as probes, primers, and markers of cancer
PT progression.
XX
PS Example 1; Page 298; 378pp; English.
XX
CC The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumour
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumour polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumour DNA or protein by incubating isolated T-
CC cells allowing them to proliferate, and administering to the patient. The
CC sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549
CC represent human ovarian tumour protein cDNA clones
XX
SQ Sequence 240 BP; 46 A; 63 C; 60 G; 66 T; 0 U; 5 Other;

Query Match 6.2%; Score 137; DB 4; Length 240;
Best Local Similarity 99.5%; Pred. No. 1.9e-57;
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 188 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 129

Qy 1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 128 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 69

Qy 1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 68 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 9

```
Qy       1575 GGGAAACA 1582
              ||||||||
Db          8 GGGAAACA 1
```

RESULT 171
AAH83671/c
ID AAH83671 standard; cDNA; 240 BP.
XX
AC AAH83671;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:1295.
XX
KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200151513-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US001575.
XX
PR 14-JAN-2000; 2000US-0176722P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA;
XX
DR WPI; 2001-425866/45.
XX
PT Novel ovarian tumor proteins, and nucleic acids encoding them, used to
PT treat and diagnose cancers, particularly ovarian cancer.
XX
PS Claim 5; Page 299-300; 338pp; English.
XX
CC AAH82377 to AAH83878 represent human ovarian tumour-associated
CC polynucleotide sequences which encode ovarian tumour proteins. The
CC ovarian tumour protein and polynucleotide sequences have cytostatic
CC activity, and can be used in gene therapy and vaccine production. The
CC ovarian tumour proteins and polynucleotides can be used to inhibit the
CC development of cancer, particularly ovarian cancer. They can also be used
CC to diagnose the onset and progression of cancer
XX
SQ Sequence 240 BP; 46 A; 63 C; 60 G; 66 T; 0 U; 5 Other;

Query Match 6.2%; Score 137; DB 5; Length 240;
Best Local Similarity 99.5%; Pred. No. 1.9e-57;
Matches 187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 129

Qy       1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        128 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 69

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 9

Qy       1575 GGGAAACA 1582
              ||||||||
Db          8 GGGAAACA 1
```

RESULT 172
ADL24638
ID ADL24638 standard; DNA; 177 BP.
XX
AC ADL24638;
XX
DT 20-MAY-2004 (first entry)
XX
DE Intestinal epithelium/peyer's patch M cell-associated DNA sequence #148.
XX
KW intestinal epithelium cell development; peyer's patch M cell development;
KW inflammatory bowel disease; glutenenteropathy; infectious disease;
KW autoimmune disease; haemolytic anaemia; rheumatoid arthritis; dermatitis;
KW Grave's disease; multiple sclerosis; allergy; asthma; diabetic mellitus;
KW immune system disorder; hypersensitivity; anaphylaxis;
KW blood group incompatibility; ds; human.
XX
OS Homo sapiens.
XX
PN WO200280852-A2.
XX
PD 17-OCT-2002.
XX
PF 04-APR-2002; 2002WO-US010873.
XX
PR 04-APR-2001; 2001US-0281416P.
XX
PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
PI Brayden DJ, Byrne D, O'mahony DJ, Evans CF, Mah SP, Lo DD;
XX
DR WPI; 2003-075470/07.
XX
PT Novel isolated or purified polypeptide encoded by genes associated with
PT intestinal epithelium or M cell development, differentiation or function,
PT useful for treating autoimmune diseases and infectious diseases.
XX
PS Claim 1; SEQ ID NO 148; 152pp; English.
XX
CC The invention comprises DNA sequences which are associated with
CC intestinal epithelium and peyer's patch M cells. The DNA sequences of the
CC invention are useful for assessing, modifying, modulating or regulating
CC intestinal epithelium or M cell development. The DNA sequences of the
CC invention are also useful in the treatment of: inflammatory bowel
CC disease, glutenenteropathy, infectious diseases, autoimmune diseases
CC (e.g. haemolytic anaemia, rheumatoid arthritis, dermatitis, Grave's
CC disease, multiple sclerosis, allergy, asthma and diabetic mellitus),
CC diseases or disorders of the immune system, hypersensitivity,
CC anaphylaxis, and blood group incompatibility. The present nucleic acid
CC represents an intestinal epithelium/peyer's patch M cell-associated DNA
CC sequence of the invention. NOTE: The present sequence is not shown in the
CC specification, but has been retrieved from the WIPO website.
XX
SQ Sequence 177 BP; 45 A; 41 C; 39 G; 52 T; 0 U; 0 Other;

Query Match 6.2%; Score 135; DB 10; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.9e-56;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          36 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 95

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          96 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 155

Qy       2176 AAAGTTGTTGCATTC 2190
              ||||||||||||||||
Db         156 AAAGTTGTTGCATTC 170
```

RESULT 173
ACD92298/c
ID ACD92298 standard; cDNA; 540 BP.
XX

AC ACD92298;
XX
DT 23-SEP-2003 (first entry)
XX
DE Human colon cancer cell expressed cDNA #710.
XX
KW Open reading frame detection; genome sequencing; colon cancer;
KW breast cancer; population genome analysis; genetic shift; cancer;
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
KW agriculture; food crop genome; resistance gene; retrovirus;
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN US2002155438-A1.
XX
PD 24-OCT-2002.
XX
PF 27-SEP-1999; 99US-00406117.
XX
PR 20-NOV-1998; 98US-00196716.
XX
PA (SIMP/) SIMPSON A J G.
PA (NETO/) NETO E D.
PA (BREN/) BRENTANI R R.
XX
PI Simpson AJG, Neto ED, Brentani RR;
XX
DR WPI; 2003-182626/18.
XX
PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.
XX
PS Example 9; Page 108; 959pp; English.
XX
CC The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA , sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame; for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a foetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX
SQ Sequence 540 BP; 115 A; 126 C; 122 G; 145 T; 0 U; 32 Other;

Query Match 6.1%; Score 134; DB 10; Length 540;
Best Local Similarity 100.0%; Pred. No. 5.7e-56;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 310 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 251

Qy 1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 250 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 191

Qy 1206 GGGATGCAGATCTT 1219
||||||||||||||
Db 190 GGGATGCAGATCTT 177

RESULT 174
ADX55963
ID ADX55963 standard; DNA; 133 BP.
XX
AC ADX55963;
XX
DT 05-MAY-2005 (first entry)
XX
DE Human ubiquitin C PCR amplicon.
XX
KW chordin like homolog; CHL2; osteopathic; gynecological; neuroprotective;
KW antiinflammatory; cytostatic; bone injury; Female genital tract disease;
KW female sexual dysfunction; miscarriage; tumor;
KW hormone metabolism disorder; cardiovascular disease;
KW neurodegenerative disease; inflammation; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2005014817-A2.
XX
PD 17-FEB-2005.
XX
PF 08-AUG-2004; 2004WO-IL000735.
XX
PR 07-AUG-2003; 2003US-0492971P.
XX
PA (COMP-) COMPUGEN LTD.
PA (OREN/) OREN A.
XX
PI Rotman G, Toporoik A, Biton S, Savitzky K, Bernstein J;
XX
DR WPI; 2005-163249/17.
XX
PT New nucleic acid sequence, useful in preparing a composition for
PT diagnosing or treating e.g., bone injuries, tumors of the lung, uterus,
PT breast or prostate, or cardiovascular, neurodegenerative or inflammatory
PT diseases.
XX
PS Example 11b; SEQ ID NO 67; 229pp; English.
XX
CC The invention relates to novel isolated chordin like homolog (CHL2)
CC nucleic acid sequences. The invention further comprises: a chordin like
CC homolog amino acid sequence; an antibody capable of specifically binding
CC to at least one epitope of the amino acid sequence; an expression vector
CC comprising the nucleic acid sequence and a control element for the
CC expression of the nucleic acid sequence in a suitable host; a host cell
CC transfected by the expression vector; a pharmaceutical composition
CC comprising a carrier and as an active ingredient, the expression vector,
CC nucleic acid sequence, amino acid sequence or antibody; a method for
CC treating a disease; a method for detecting a Chordin like homologue
CC (CHL2) nucleic acid sequence in a biological sample; a method for
CC detecting antibodies in a biological sample; a biomarker for detecting
CC prostate cancer, comprising hCHL2 variant X having a 1515-nt sequence
CC (ADX55906) or its fragment; a primer pair for use in detecting the

CC biomarker; an amplicon obtained through the use of the primer pair; a
CC nucleic acid molecule capable of selectively hybridizing to hCHL2
CC variants IV, V, VI, VII, VIII, IX or X; a biomarker for detecting cancers
CC comprising the sequence of exon 2a (ADX55960); a kit for detecting
CC prostate cancer; a kit for detecting prostate cancer of cells; a method
CC for detecting prostate cancer in cells suspected of being cancerous; a
CC biomarker for detecting prostate cancer; a biomarker for detecting breast
CC or lung cancer; a method for detecting lung or breast cancer in cells
CC suspected of being cancerous; and a method for detecting a CHL2-product
CC in a biological sample. The chordin like homolog proteins, encoding
CC nucleic acids and compositions have osteopathic, gynecological,
CC neuroprotective, antiinflammatory, and cytostatic activities. The chordin
CC like homolog nucleic acid sequence is useful in preparing a composition
CC for diagnosing or treating bone injuries. It is also useful in treating
CC diseases involved with the female reproductive tract, diseases or
CC disorders involved with abnormal sexual differentiation, recurrent
CC miscarriages, tumors of the lung, uterus, breast or prostate, diseases
CC involving sexual hormone abnormalities, cardiovascular disorders,
CC neuronal diseases of the CNS, neurodegenerative diseases and diseases
CC involving non-normal developments of neurons, inflammatory diseases, or
CC diseases of the gut. This polynucleotide sequence represents a human
CC ubiquitin C house-keeping gene PCR amplicon used in comparison to human
CC CHL2 of the invention.
XX
SQ Sequence 133 BP; 30 A; 31 C; 36 G; 36 T; 0 U; 0 Other;

Query Match 6.1%; Score 133; DB 14; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy 80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy 140 GAATGTCAAGGCA 152
|||||||||||||
Db 121 GAATGTCAAGGCA 133

RESULT 175
AEB53761
ID AEB53761 standard; DNA; 133 BP.
XX
AC AEB53761;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human ubiquitin amplicon SEQ ID NO 33.
XX
KW cytostatic; neuroprotective; nootropic; antidepressant; anti-HIV;
KW antiallergic; antianemic; antiasthmatic; antiarteriosclerotic;
KW dermatological; antidiabetic; osteopathic; antipsoriatic;
KW antiinflammatory; antiulcer; gastrointestinal-gen.; gene therapy;
KW diagnosis; prognosis; therapeutic; pharmaceutical; cancer; neoplasm;
KW neurological disease; nootropic; neuroleptic; antidepressant;
KW psychiatric disorder; autoimmune disease; immunosuppressive; anti-hiv;
KW antiallergic; antianemic; hematological disease; antiasthmatic;
KW respiratory disease; antiarteriosclerotic; cardiovascular disease;
KW metabolic disorder; dermatological; antidiabetic; metabolic disorder;
KW antigout; neuroprotective; osteopathic; degeneration; endocrine disease;
KW musculoskeletal disease; antipsoriatic; dermatological disease;
KW antiinflammatory; antiulcer; gastrointestinal-gen.;
KW gastrointestinal disease; immune disorder; inflammation; ubiquitin; ds.
XX
OS Homo sapiens.
XX
PN WO2005068618-A1.
XX
PD 28-JUL-2005.
XX
PF 13-JAN-2005; 2005WO-IL000047.
XX
PR 13-JAN-2004; 2004US-0535904P.
PR 19-MAY-2004; 2004US-0572122P.
XX
PA (COMP-) COMPUGEN LTD.
XX
PI Sella-Tavor O, Rotman G, Pollock S, Diber A, Walach S;
PI Sameach-Greenwald S;
XX
DR WPI; 2005-555014/56.
XX
PT New isolated UbcH10 polynucleotides and polypeptides, useful for
PT diagnosing or treating UbcH10-related diseases, e.g. ovarian cancer, lung
PT cancer, Alzheimer's disease, amnesia, allergies, anemia, asthma,
PT atherosclerosis or diabetes.
XX
PS Example 1; SEQ ID NO 33; 165pp; English.
XX
CC The invention describes an isolated polynucleotide comprising fully
CC defined 696-935 base pair sequences (SEQ ID NO. 1-3) given in the
CC specification, and a nucleic acid sequence encoding a UbcH10 polypeptide
CC having at least a portion of an amino acid sequence at least 70%
CC homologous to fully defined 50 amino acid sequence (SEQ ID NO. 7) given
CC in the specification. Also described are: an isolated polynucleotide
CC encoding UbcH10; an antibody or an antibody fragment capable of
CC specifically binding the isolated polypeptide above; an antibody or an
CC antibody fragment capable of specifically binding a UbcH10 polypeptide;
CC an oligonucleotide specifically hybridizable with a nucleic acid sequence
CC encoding a UbcH10 polypeptide including at least a portion of an amino
CC acid sequence at least 70% homologous to SEQ ID NO. 7 or 8, as determined
CC using the BlastP software of the NCBI using default parameters;
CC diagnosing predisposition to, or presence of a UbcH10-related disease in
CC a subject; and a kit for diagnosing UbcH10-related disease or a
CC predisposition to it in a subject, the kit comprising at least one
CC reagent capable of detecting overexpression of the UbcH10 polypeptide
CC above. Also disclosed are: a display library comprising display vehicles
CC displaying at least 6 consecutive amino acids derived from a UbcH10
CC polypeptide; a pharmaceutical composition comprising a UbcH10 polypeptide
CC ; and treating UbcH10-related disease in a subject. The polynucleotides
CC and polypeptides are useful for diagnosing and treating UbcH10-related
CC diseases. Diseases include cancer, e.g. ovarian cancer, lung cancer,
CC colorectal cancer, or oral cancer; neuronal diseases, e.g. Alzheimer's
CC disease, amnesia, bipolar disorder, or depression; and autoimmune
CC disorders, e.g. AIDS, allergies, anemia, asthma, atherosclerosis, contact
CC dermatitis, diabetes, gout, multiple sclerosis, osteoporosis, psoriasis,
CC or ulcerative colitis. This sequence represents an amplicon of human
CC ubiquitin, a house keeping gene used in an assay to measure UbcH10
CC expression.
XX
SQ Sequence 133 BP; 30 A; 31 C; 36 G; 36 T; 0 U; 0 Other;

Query Match 6.1%; Score 133; DB 14; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy 80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy 140 GAATGTCAAGGCA 152
|||||||||||||
Db 121 GAATGTCAAGGCA 133

RESULT 176
AEC04911
ID AEC04911 standard; DNA; 133 BP.

```
XX
AC   AEC04911;
XX
DT   20-OCT-2005  (first entry)
XX
DE   Human UBC amplicon.
XX
KW   Cytostatic; Gene therapy; diagnosis; breast tumor; endocrine disease;
KW   gynecology and obstetrics; neoplasm; tumor marker; ds.
XX
OS   Homo sapiens.
XX
PN   WO2005072050-A2.
XX
PD   11-AUG-2005.
XX
PF   27-JAN-2005; 2005WO-IB000433.
XX
PR   27-JAN-2004; 2004US-0539128P.
PR   27-JAN-2004; 2004US-0539129P.
PR   22-OCT-2004; 2004US-0620656P.
PR   22-OCT-2004; 2004US-0620853P.
PR   22-OCT-2004; 2004US-0620874P.
PR   22-OCT-2004; 2004US-0620916P.
PR   22-OCT-2004; 2004US-0620917P.
PR   22-OCT-2004; 2004US-0620918P.
PR   22-OCT-2004; 2004US-0620924P.
PR   22-OCT-2004; 2004US-0620974P.
PR   22-OCT-2004; 2004US-0620975P.
PR   22-OCT-2004; 2004US-0621004P.
PR   25-OCT-2004; 2004US-0621131P.
PR   17-NOV-2004;   55US-00043842.
PR   17-NOV-2004; 2004US-0620123P.
PR   17-NOV-2004; 2004US-0628101P.
PR   17-NOV-2004; 2004US-0628111P.
PR   17-NOV-2004; 2004US-0628112P.
PR   17-NOV-2004; 2004US-0628134P.
PR   17-NOV-2004; 2004US-0628145P.
PR   17-NOV-2004; 2004US-0628156P.
PR   17-NOV-2004; 2004US-0628167P.
PR   17-NOV-2004; 2004US-0628178P.
PR   17-NOV-2004; 2004US-0628231P.
PR   17-NOV-2004; 2004US-0628251P.
PR   27-JAN-2005; 2005US-00043842.
XX
PA   (COMP-) COMPUGEN USA INC.
XX
PI   Toporik A,  Dahary D,  Sorek R,  Pollock S,  Levine Z,  Akiva P;
PI   Diber A,  Novik A,  Sella-Tavor O,  Ayalon-Soffer M,  Walach S;
PI   Sameah-Greenwald S,  Shemesh R,  Keren N,  Shklar M;
XX
DR   WPI; 2005-555592/56.
XX
PT   New human nucleic acid and polypeptide sequences useful for screening,
PT   diagnosing or treating breast cancer.
XX
PS   Disclosure; Page 193; 1586pp; English.
XX
CC   The invention relates to an isolated human polynucleotide. The
CC   composition and methods are useful for screening, diagnosing or treating
CC   breast cancer. These may also be used in drug screening or in monitoring
CC   disease progression and/or treatment efficacy of breast cancer. The
CC   present sequence represents a human housekeeping gene amplicon.
XX
SQ   Sequence 133 BP; 30 A; 31 C; 36 G; 36 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 133;  DB 14;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCA 152
              |||||||||||||
Db        121 GAATGTCAAGGCA 133


RESULT 177
AEB63482
ID   AEB63482 standard; DNA; 133 BP.
XX
AC   AEB63482;
XX
DT   20-OCT-2005  (first entry)
XX
DE   Human ubiquitin probe SEQ ID NO 38.
XX
KW   cardiant; antianginal; cardiovascular-gen.; diagnostic;
KW   protein purification; myocardial infarction; angina; antiangina;
KW   myocardial disease; myocarditis; cardiant; congestive heart failure;
KW   cardiovascular-gen; cardiovascular disease; ubiquitin; probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO2005072055-A2.
XX
PD   11-AUG-2005.
XX
PF   27-JAN-2005; 2005WO-IB000995.
XX
PR   27-JAN-2004; 2004US-0539127P.
PR   15-JUL-2004; 2004US-0587851P.
PR   07-SEP-2004; 2004US-0607246P.
PR   25-OCT-2004; 2004US-0622320P.
PR   27-OCT-2004; 2004US-0622016P.
PR   17-NOV-2004; 2004US-0628190P.
PR   27-JAN-2005; 2005US-00043590.
XX
PA   (COMP-) COMPUGEN USA INC.
XX
PI   Cojocaru GS,  Diber A,  Novik A,  Pollock S,  Levine Z,  Cohen Y;
PI   Ayalon-Soffer M;
XX
DR   WPI; 2005-542336/55.
XX
PT   New isolated polynucleotide encoding Brain Natriuretic Peptide (BNP),
PT   useful for diagnosing or treating cardiac diseases e.g. myocardial
PT   infarct, acute coronary syndrome, angina pectoris, cardiomyopathy, or
PT   myocarditis.
XX
PS   Example 3; SEQ ID NO 38; 127pp; English.
XX
CC   The invention describes an isolated polynucleotide comprising a nucleic
CC   acid sequence selected from HUMNATPEP_PEA 1 _T1, HUMNATPEP_PEA 1 _T2,
CC   HUMNATPEP_PEA 1 _T3, or HUMNATPEP_PEA 1 _T4. Specifically claimed are
CC   polynucleotides encoding Brain Natriuretic Peptide (BNP). The
CC   polynucleotide is selected from HUMNATPEP_PEA 1 _T1, HUMNATPEP_PEA 1 _T2,
CC   HUMNATPEP_PEA 1 _T3, or HUMNATPEP_PEA 1 _T4. The polynucleotides and
CC   polypeptides are useful for the diagnosis or treatment of BNP-related
CC   diseases including cardiac diseases, e.g. myocardial infarct, acute
CC   coronary syndrome, angina pectoris, cardiomyopathy, myocarditis, or
CC   congestive heart failure. This sequence represents a ubiquitin probe used
CC   as a control in the expression analysis of natriuretic peptides B
CC   precursor variants of the invention.
XX
SQ   Sequence 133 BP; 30 A; 31 C; 36 G; 36 T; 0 U; 0 Other;

  Query Match             6.1%;  Score 133;  DB 14;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-55;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy          80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy         140 GAATGTCAAGGCA 152
               |||||||||||||
Db         121 GAATGTCAAGGCA 133
```

RESULT 178
AEB70674
ID AEB70674 standard; DNA; 133 BP.
XX
AC AEB70674;
XX
DT 20-OCT-2005 (first entry)
XX
DE Endometriosis marker associated ubiquitin amplicon.
XX
KW gynecological; gene therapy; diagnosis; selectable marker; endometriosis;
KW gynecological; genitourinary disease; gynecology and obstetrics;
KW ubiquitin; ds.
XX
OS Unidentified.
XX
PN WO2005072049-A2.
XX
PD 11-AUG-2005.
XX
PF 27-JAN-2005; 2005WO-IB001188.
XX
PR 27-JAN-2004; 2004US-0539128P.
PR 27-JAN-2004; 2004US-0539129P.
PR 22-OCT-2004; 2004US-0621004P.
PR 17-NOV-2004; 2004US-0628145P.
PR 17-NOV-2004; 2004US-0628178P.
PR 17-NOV-2004; 2004US-0628230P.
PR 27-JAN-2005; 55US-00621004.
XX
PA (COMP-) COMPUGEN USA INC.
XX
PI Cohen Y, Pollock S, Novik A, Diber A;
XX
DR WPI; 2005-555591/56.
XX
PT New mammalian nucleic acid and polypeptide sequences useful for
PT screening, diagnosing or treating endometriosis, or in monitoring disease
PT progression or efficacy of treatment for endometriosis.
XX
PS Disclosure; SEQ ID NO 445; 804pp; English.
XX
CC The invention describes a nucleic acid sequence comprising any of the 7
CC nucleotide sequences having 29-1404 bp fully defined in the specification
CC (SEQ ID NOS: 1-7). Also described are: an amino acid sequence comprising
CC 65 amino acids fully defined in the specification (SEQ ID NO: 9); an
CC isolated chimeric polypeptide encoding for S71513_P2, comprising a first
CC amino acid sequence being at least 90% homologous to Met-Lys-Val-Ser-Ala-
CC Ala-Leu-Leu-Cys-Leu-Leu-Leu-Ile-Ala-Ala-Thr-Phe-Ile-Pro-Gln-Gly-Leu-Ala-
CC Gln-Pro-Asp-Ala-Ile-Asn-Ala-Pro-Val-Thr-Cys-Cys-Tyr-Asn-Phe-Thr-Asn-Arg-
CC Lys-Ile-Ser-Val-Gln-Arg-Leu-Ala-Ser-Tyr-Arg-Arg-Ile-Thr-Ser-Ser-Lys-Cys-
CC Pro-Lys-Glu-Ala-Val corresponding to amino acids 1-64 of SY02_HUMAN,
CC which also corresponds to amino acids 1-64 of S71513_P2, and a second
CC amino acid sequence comprising a polypeptide having the sequence M
CC corresponding to amino acid 65 of S71513_P2, where the first amino acid
CC sequence and second amino acid sequence are contiguous and in a
CC sequential order; an antibody capable of specifically binding to an
CC epitope of the amino acid sequence mentioned above; a kit for detecting
CC endometriosis, comprising a kit detecting overexpression of the splice
CC variant or a kit comprising the above antibody; methods for detecting,
CC diagnosing or screening for endometriosis; a biomarker capable of
CC detecting endometriosis, comprising the above nucleic acid or amino acid
CC sequences, or their fragments; a method for monitoring disease
CC progression and/or treatment efficacy and/or relapse of endometriosis;
CC and a method of selecting a therapy for endometriosis. The composition
CC and methods are useful for screening, diagnosing or treating
CC endometriosis. These may also be used in monitoring disease progression
CC and the efficacy of treatment for endometriosis. This sequence represents
CC an ubiquitin amplicon associated with the identification of candidate
CC markers for endometriosis.
XX
SQ Sequence 133 BP; 30 A; 31 C; 36 G; 36 T; 0 U; 0 Other;

Query Match 6.1%; Score 133; DB 14; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy          80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy         140 GAATGTCAAGGCA 152
               |||||||||||||
Db         121 GAATGTCAAGGCA 133
```

RESULT 179
ADR12568
ID ADR12568 standard; DNA; 228 BP.
XX
AC ADR12568;
XX
DT 04-NOV-2004 (first entry)
XX
DE Gene vaccine nucleic acid #18.
XX
KW ds; gene; antimalarial; antitubercular; tuberculostatic; anti-HIV;
KW antibacterial; haemostatic; protozoacide; antiinflammatory;
KW neuroprotective; virucide; gene vaccine; ubiquitin; antigen;
KW intracellular parasite; protozoan parasite infection; malaria;
KW tuberculosis; toxoplasmosis; trypanosomiasis; AIDS;
KW cytomegalovirus infection; achlamydia disease; Rickettsia; leishmaniasis;
KW Ebola hemorrhagic fever; Trypanosoma; Chagas disease;
KW Japanese encephalitis; influenza; rubeola; dengue virus; poliomyelitis;
KW Herpes virus; severe acute respiratory syndrome.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .228
FT /*tag= a
XX
PN WO2004067040-A1.
XX
PD 12-AUG-2004.
XX
PF 30-JAN-2004; 2004WO-JP000975.
XX
PR 31-JAN-2003; 2003JP-00023507.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Ishii K;
XX
DR WPI; 2004-594036/57.
DR P-PSDB; ADR12569.
XX
PT Gene vaccine having nucleic acids encoding ubiquitin and antigen protein

PT of intracellular parasite, useful for treating parasite infections such
PT as malaria, tuberculosis, toxoplasmosis.
XX
PS Disclosure; SEQ ID NO 35; 266pp; Japanese.
XX
CC A gene vaccine (I) comprises nucleic acid sequence encoding ubiquitin,
CC and a nucleic acid sequence encoding the antigen protein of an
CC intracellular parasite containing T-cell target sequence. (I) is useful
CC for preventing or treating the disease resulting from intracellular
CC protozoan parasite infection such as malaria, tuberculosis,
CC toxoplasmosis, trypanosomiasis, AIDS, cytomegalovirus infection,
CC achlamydia disease, infections caused by Rickettsia , leishmaniasis,
CC Ebola hemorrhagic fever, Trypanosoma infections, Chagas disease, Japanese
CC encephalitis, influenza, rubeola and dengue viral infections,
CC poliomyelitis, Herpes virus (alpha) infections, or severe acute
CC respiratory syndrome. This sequence represent a nucleic acid used in the
CC method of the invention.
XX
SQ Sequence 228 BP; 63 A; 62 C; 62 G; 41 T; 0 U; 0 Other;

Query Match 5.7%; Score 126; DB 13; Length 228;
Best Local Similarity 99.1%; Pred. No. 5.5e-52;
Matches 226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      297 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGT 356
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGT 60

Qy      357 GACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 416
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       61 GACACCATCGAGAACGTCAAGGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 120

Qy      417 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 476
            |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      121 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGGACCCTGTCTGACTACAAC 180

Qy      477 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 524
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 228
```

RESULT 180
ADS87021
ID ADS87021 standard; DNA; 228 BP.
XX
AC ADS87021;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genetic vaccine/ubiquitin (Ub)-related DNA - SEQ ID 37.
XX
KW vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; human; ds; gene.
XX
OS Homo sapiens.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
DR P-PSDB; ADS87022.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Disclosure; SEQ ID NO 37; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a human genetic vaccine/ubiquitin (Ub)-related DNA of the
CC invention.
XX
SQ Sequence 228 BP; 63 A; 62 C; 62 G; 41 T; 0 U; 0 Other;

Query Match 5.7%; Score 126; DB 13; Length 228;
Best Local Similarity 99.1%; Pred. No. 5.5e-52;
Matches 226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      297 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGT 356
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGT 60

Qy      357 GACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 416
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       61 GACACCATCGAGAACGTCAAGGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 120

Qy      417 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGGACCCTGTCTGACTACAAC 476
            |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      121 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGGACCCTGTCTGACTACAAC 180

Qy      477 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 524
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 228
```

RESULT 181
AAD00389
ID AAD00389 standard; DNA; 369 BP.
XX
AC AAD00389;
XX
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX
DE Ubiquitin-E. coli beta-galactosidase fusion protein encoding DNA UBILacI.
XX
KW Ubiquitin; beta-galactosidase; LacI; fusion protein; UBILacI; human;
KW Mucin 1; MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW immune response; cytostatic; vaccine; ds.
XX
OS Homo sapiens.
OS Escherichia coli.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1. .369
FT /*tag= a
FT /product= "Ubiquitin-partial beta-galactosidase fusion"
FT /note= "Does not include stop codon"
FT /partial
XX
PN WO200025827-A2.
XX
PD 11-MAY-2000.
XX
PF 18-OCT-1999; 99WO-EP007874.
XX
PR 30-OCT-1998; 98IT-MI002330.
XX
PA (MENA) MENARINI RICERCHE SPA.

```
XX
PI   Parente D,  Di Massimo AM,  De Santis R;
XX
DR   WPI; 2000-365410/31.
DR   P-PSDB; AAY71025.
XX
PT   Composition containing one or more DNA molecules encoding fragments of a
PT   Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT   tumor therapy.
XX
PS   Claim 17; Fig 6; 56pp; English.
XX
CC   The present sequence is a DNA termed UBILacI that encodes a fusion
CC   protein containing ubiquitin from MCF7 tumour cell line and a partial
CC   sequence of E. coli beta-galactosidase (LacI portion). This sequence is
CC   fused to the 5' terminus of DNA molecules encoding fragments of human
CC   Mucin 1 (MUC-1), an antigenic protein which is overexpressed in tumour
CC   cells. The DNA molecules encoding MUC-1 fragments are cloned into pMRS30
CC   expression vector and used in pharmaceutical compositions e.g. vaccine
CC   for inducing an antigen-specific anti-tumour immune response.
CC   Compositions containing these DNA molecules are useful in anti-tumour
CC   therapy of patients affected with tumours characterised by high MUC-1
CC   expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 369 BP; 93 A; 108 C; 98 G; 70 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 126;  DB 3;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-52;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 182
AAD00390
ID   AAD00390 standard; DNA; 579 BP.
XX
AC   AAD00390;
XX
DT   15-SEP-2003  (revised)
DT   29-AUG-2000  (first entry)
XX
DE   Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #1.
XX
KW   Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW   MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW   immune response; cytostatic; vaccine; ds.
XX
OS   Homo sapiens.
OS   Escherichia coli.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .579
FT                   /*tag=  a
FT                   /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT   misc_feature    1. .369
FT                   /*tag=  b
FT                   /label= UBILacI_DNA
FT                   /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT   misc_feature    370. .579
FT                   /*tag=  c
FT                   /note= "Human MUC-1 partial DNA that corresponds to
FT                   nucleotides 136-339 of the EMBL sequence J05581 with two
FT                   stop codons"
XX
PN   WO200025827-A2.
XX
PD   11-MAY-2000.
XX
PF   18-OCT-1999;   99WO-EP007874.
XX
PR   30-OCT-1998;   98IT-MI002330.
XX
PA   (MENA ) MENARINI RICERCHE SPA.
XX
PI   Parente D,  Di Massimo AM,  De Santis R;
XX
DR   WPI; 2000-365410/31.
DR   P-PSDB; AAY71026.
XX
PT   Composition containing one or more DNA molecules encoding fragments of a
PT   Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT   tumor therapy.
XX
PS   Claim 18; Fig 7; 56pp; English.
XX
CC   The present sequence is a DNA encoding a fusion protein consisting of
CC   human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC   terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC   and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC   protein overexpressed in tumour cells. The present sequence is cloned
CC   into a pMRS30 expression vector and used in pharmaceutical composition
CC   e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC   response. Composition containing this DNA molecule is useful in anti-
CC   tumour therapy of patients affected with tumours characterised by high
CC   MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 579 BP; 145 A; 172 C; 153 G; 109 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 126;  DB 3;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-52;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 183
AAD00393
ID   AAD00393 standard; DNA; 651 BP.
XX
AC   AAD00393;
XX
DT   15-SEP-2003  (revised)
DT   29-AUG-2000  (first entry)
XX
DE   Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #4.
XX
KW   Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW   MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW   immune response; cytostatic; vaccine; ds.
XX
OS   Homo sapiens.
OS   Escherichia coli.
OS   Chimeric.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             1. .651
FT                   /*tag= a
FT                   /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT   misc_feature    1. .369
FT                   /*tag= b
FT                   /label= UBILacI_DNA
FT                   /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT   misc_feature    370. .651
FT                   /*tag= c
FT                   /note= "Human MUC-1 partial DNA that corresponds to
FT                   nucleotides 1222-1497 of the EMBL sequence J05581 with
FT                   two stop codons"
XX
PN   WO200025827-A2.
XX
PD   11-MAY-2000.
XX
PF   18-OCT-1999;   99WO-EP007874.
XX
PR   30-OCT-1998;   98IT-MI002330.
XX
PA   (MENA ) MENARINI RICERCHE SPA.
XX
PI   Parente D,  Di Massimo AM,  De Santis R;
XX
DR   WPI; 2000-365410/31.
DR   P-PSDB; AAY71029.
XX
PT   Composition containing one or more DNA molecules encoding fragments of a
PT   Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT   tumor therapy.
XX
PS   Claim 18; Fig 10; 56pp; English.
XX
CC   The present sequence is a DNA encoding a fusion protein consisting of
CC   human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC   terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC   and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC   protein overexpressed in tumour cells. The present sequence is cloned
CC   into a pMRS30 expression vector and used in pharmaceutical composition
CC   e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC   response. Composition containing this DNA molecule is useful in anti-
CC   tumour therapy of patients affected with tumours characterised by high
CC   MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 651 BP; 149 A; 195 C; 168 G; 139 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 126;  DB 3;  Length 651;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-52;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 184
AAD00391
ID   AAD00391 standard; DNA; 891 BP.
XX
AC   AAD00391;
XX
DT   15-SEP-2003  (revised)
DT   29-AUG-2000  (first entry)
XX
DE   Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #2.
XX
KW   Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW   MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW   immune response; cytostatic; vaccine; ds.
XX
OS   Homo sapiens.
OS   Escherichia coli.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .891
FT                   /*tag= a
FT                   /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT   misc_feature    1. .369
FT                   /*tag= b
FT                   /label= UBILacI_DNA
FT                   /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT   misc_feature    370. .891
FT                   /*tag= c
FT                   /note= "Human MUC-1 partial DNA that corresponds to
FT                   nucleotides 205-720 of the EMBL sequence J05581 with two
FT                   stop codons"
XX
PN   WO200025827-A2.
XX
PD   11-MAY-2000.
XX
PF   18-OCT-1999;   99WO-EP007874.
XX
PR   30-OCT-1998;   98IT-MI002330.
XX
PA   (MENA ) MENARINI RICERCHE SPA.
XX
PI   Parente D,  Di Massimo AM,  De Santis R;
XX
DR   WPI; 2000-365410/31.
DR   P-PSDB; AAY71027.
XX
PT   Composition containing one or more DNA molecules encoding fragments of a
PT   Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT   tumor therapy.
XX
PS   Claim 18; Fig 8; 56pp; English.
XX
CC   The present sequence is a DNA encoding a fusion protein consisting of
CC   human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC   terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC   and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC   protein overexpressed in tumour cells. The present sequence is cloned
CC   into a pMRS30 expression vector and used in pharmaceutical composition
CC   e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC   response. Composition containing this DNA molecule is useful in anti-
CC   tumour therapy of patients affected with tumours characterised by high
CC   MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 891 BP; 203 A; 307 C; 225 G; 156 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 126;  DB 3;  Length 891;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-52;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
```

```
Db        121 CAGAGG 126

RESULT 185
AAD00392
ID   AAD00392 standard; DNA; 1020 BP.
XX
AC   AAD00392;
XX
DT   15-SEP-2003  (revised)
DT   29-AUG-2000  (first entry)
XX
DE   Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #3.
XX
KW   Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW   MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW   immune response; cytostatic; vaccine; ds.
XX
OS   Homo sapiens.
OS   Escherichia coli.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1020
FT                   /*tag=  a
FT                   /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT   misc_feature    1. .369
FT                   /*tag=  b
FT                   /label= UBILacI_DNA
FT                   /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT   misc_feature    370. .1020
FT                   /*tag=  c
FT                   /note= "Human MUC-1 partial DNA that corresponds to
FT                   nucleotides 631-1275 of the EMBL sequence J05581 with two
FT                   stop codons"
XX
PN   WO200025827-A2.
XX
PD   11-MAY-2000.
XX
PF   18-OCT-1999;   99WO-EP007874.
XX
PR   30-OCT-1998;   98IT-MI002330.
XX
PA   (MENA ) MENARINI RICERCHE SPA.
XX
PI   Parente D,  Di Massimo AM,  De Santis R;
XX
DR   WPI; 2000-365410/31.
DR   P-PSDB; AAY71028.
XX
PT   Composition containing one or more DNA molecules encoding fragments of a
PT   Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT   tumor therapy.
XX
PS   Claim 18; Fig 9; 56pp; English.
XX
CC   The present sequence is a DNA encoding a fusion protein consisting of
CC   human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC   terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC   and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC   protein overexpressed in tumour cells. The present sequence is cloned
CC   into a pMRS30 expression vector and used in pharmaceutical composition
CC   e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC   response. Composition containing this DNA molecule is useful in anti-
CC   tumour therapy of patients affected with tumours characterised by high
CC   MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 1020 BP; 236 A; 303 C; 237 G; 244 T; 0 U; 0 Other;

 Query Match             5.7%;  Score 126;  DB 3;  Length 1020;
 Best Local Similarity   100.0%;  Pred. No. 5.3e-52;
 Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126

RESULT 186
AAD00394
ID   AAD00394 standard; DNA; 1737 BP.
XX
AC   AAD00394;
XX
DT   15-SEP-2003  (revised)
DT   29-AUG-2000  (first entry)
XX
DE   Ubiquitin-E. coli LacI-human Mucin 1 fusion protein encoding DNA #5.
XX
KW   Ubiquitin; LacI; beta-galactosidase; fusion protein; human; Mucin 1;
KW   MUC-1; tumour; pMRS30 expression vector; anti-tumour; therapy;
KW   immune response; cytostatic; vaccine; ds.
XX
OS   Homo sapiens.
OS   Escherichia coli.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1737
FT                   /*tag=  a
FT                   /product= "Ubiquitin-LacI-MUC-1 fusion protein"
FT   misc_feature    1. .369
FT                   /*tag=  b
FT                   /label= UBILacI_DNA
FT                   /note= "Includes ubiquitin-E. coli LacI fusion DNA"
FT   misc_feature    370. .1737
FT                   /*tag=  c
FT                   /note= "Human MUC-1 partial DNA that corresponds to
FT                   nucleotides 136-1497 of the EMBL sequence J05581 with two
FT                   stop codons"
XX
PN   WO200025827-A2.
XX
PD   11-MAY-2000.
XX
PF   18-OCT-1999;   99WO-EP007874.
XX
PR   30-OCT-1998;   98IT-MI002330.
XX
PA   (MENA ) MENARINI RICERCHE SPA.
XX
PI   Parente D,  Di Massimo AM,  De Santis R;
XX
DR   WPI; 2000-365410/31.
DR   P-PSDB; AAY71030.
XX
PT   Composition containing one or more DNA molecules encoding fragments of a
PT   Mucin 1 (MUC-1) protein overexpressed in tumor cells, useful in anti-
PT   tumor therapy.
XX
PS   Claim 18; Fig 11; 56pp; English.
XX
CC   The present sequence is a DNA encoding a fusion protein consisting of
CC   human Mucin 1 (MUC-1) fragment fused to UBILacI sequence at the N-
CC   terminus. The UBILacI sequence consists of ubiquitin from MCF7 cell line
CC   and a portion of E. coli beta-galactosidase (LacI). MUC-1 is an antigenic
CC   protein overexpressed in tumour cells. The present sequence is cloned
```

```
CC   into a pMRS30 expression vector and used in pharmaceutical composition
CC   e.g. vaccine for inducing an antigen-specific anti-tumour immune
CC   response. Composition containing this DNA molecule is useful in anti-
CC   tumour therapy of patients affected with tumours characterised by high
CC   MUC-1 expression. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 1737 BP; 392 A; 558 C; 418 G; 369 T; 0 U; 0 Other;

  Query Match             5.7%;  Score 126;  DB 3;  Length 1737;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-52;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1557 CAGAGG 1562
             ||||||
Db        121 CAGAGG 126


RESULT 187
AAH69111
ID   AAH69111 standard; cDNA; 255 BP.
XX
AC   AAH69111;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 385.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 170; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 255 BP; 70 A; 68 C; 63 G; 44 T; 0 U; 10 Other;

  Query Match             5.7%;  Score 125;  DB 4;  Length 255;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-51;
  Matches  125;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 79

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 139

Qy        870 CAGCA 874
              |||||
Db        140 CAGCA 144


RESULT 188
AAA69789/c
ID   AAA69789 standard; cDNA; 636 BP.
XX
AC   AAA69789;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Human ovarian carcinoma antigen polynucleotide SEQ ID NO:99.
XX
KW   Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KW   tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200036107-A2.
XX
PD   22-JUN-2000.
XX
PF   17-DEC-1999;   99WO-US030270.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Frudakis TN;
XX
DR   WPI; 2000-431589/37.
XX
PT   Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT   encoding it, useful for the diagnosis, prevention and treatment of
PT   cancer, preferably ovarian cancer.
XX
PS   Claim 18; Fig 15; 299pp; English.
XX
CC   The present invention describes an isolated polypeptide comprising an
CC   immunogenic portion of an ovarian carcinoma protein (or its variants).
CC   Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC   cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC   carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC   for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC   cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557 represent human
CC   ovarian carcinoma polynucleotides and proteins used in the
CC   exemplification of the present invention
XX
SQ   Sequence 636 BP; 137 A; 158 C; 158 G; 165 T; 0 U; 18 Other;

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-50;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
```

```
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db         459 GAA 457


RESULT 189
ABN72683/c
ID   ABN72683 standard; DNA; 636 BP.
XX
AC   ABN72683;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Ovarian carcinoma antigen polynucleotide #99.
XX
KW   Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS   Homo sapiens.
XX
PN   WO200206317-A2.
XX
PD   24-JAN-2002.
XX
PF   17-JUL-2001; 2001WO-US022635.
XX
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
PR   04-APR-2001; 2001US-00827271.
PR   18-JUN-2001; 2001US-00884441.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Fling SP,  Retter MW,  Fanger GR;
PI   Reed SG,  Vedvick TS,  Carter D,  Hill P,  Albone E;
XX
DR   WPI; 2002-164781/21.
XX
PT   Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT   protein or its variants, useful for stimulating an immune response in a
PT   patient and treating ovarian cancer.
XX
PS   Example 1; Page 132; 408pp; English.
XX
CC   This invention relates to polypeptides comprising an immunogenic portion
CC   of an ovarian carcinoma protein which acts as an immunostimulant and is
CC   cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC   proteins, T cell populations and antigen presenting cells that express
CC   the polypeptides are useful for stimulating an immune response in a
CC   patient and treating ovarian cancer. This sequence represents DNA related
CC   to the invention
XX
SQ   Sequence 636 BP; 137 A; 158 C; 158 G; 165 T; 0 U; 18 Other;

  Query Match             5.6%;  Score 123;  DB 6;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-50;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db         459 GAA 457


RESULT 190
ADA08848/c
ID   ADA08848 standard; DNA; 636 BP.
XX
AC   ADA08848;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human ovarian carcinoma antigen polynucleotide #99.
XX
KW   ds; human; gene therapy; ovarian cancer; cancer.
XX
OS   Homo sapiens.
XX
PN   US2003091580-A1.
XX
PD   15-MAY-2003.
XX
PF   17-JUL-2001; 2001US-00907969.
XX
PR   18-JUN-2001; 2001US-00884441.
XX
PA   (MITC/) MITCHAM J L.
PA   (KING/) KING G E.
PA   (ALGA/) ALGATE P A.
PA   (FLIN/) FLING S P.
PA   (RETT/) RETTER M W.
PA   (FANG/) FANGER G R.
PA   (REED/) REED S G.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (HILL/) HILL P.
PA   (ALBO/) ALBONE E.
XX
PI   Mitcham JL,  King GE,  Algate PA,  Fling SP,  Retter MW,  Fanger GR;
PI   Reed SG,  Vedvick TS,  Carter D,  Hill P,  Albone E;
XX
DR   WPI; 2003-532352/50.
XX
PT   New isolated O772P polypeptides and polynucleotides, useful in gene
PT   therapy, particularly for treating or diagnosing cancer, e.g. ovarian
PT   cancer.
XX
PS   Example 1; Fig 15; 371pp; English.
XX
CC   The invention relates to an isolated O772P polypeptide, which has the
CC   structure fully defined in the specification. The composition containing
CC   the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
CC   or antigen presenting cells are useful for stimulating an immune response
CC   and treating ovarian cancer. Detecting the presence of the
CC   polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
CC   carcinoma cDNAs and protein cDNAs were identified using microarray
CC   technology. The present sequence represents a human ovarian carcinoma
CC   antigen polynucleotide.
XX
SQ   Sequence 636 BP; 137 A; 158 C; 158 G; 165 T; 0 U; 18 Other;

  Query Match             5.6%;  Score 123;  DB 9;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-50;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460
```

```
Qy       1508 GAA 1510
              |||
Db        459 GAA 457


RESULT 191
ADF08595/c
ID   ADF08595 standard; cDNA; 636 BP.
XX
AC   ADF08595;
XX
DT   12-FEB-2004  (first entry)
XX
DE   cDNA encoding secreted ovarian carcinoma antigen seqid 99.
XX
KW   gene therapy; protein therapy; vaccine; antibody inhibition;
KW   breast cancer; restorative therapy; diagnostic agent; immunoassay;
KW   secreted ovarian carcinoma antigen; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003124140-A1.
XX
PD   03-JUL-2003.
XX
PF   17-JUL-2002; 2002US-00198053.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
PR   04-APR-2001; 2001US-00827271.
PR   18-JUN-2001; 2001US-00884441.
PR   17-JUL-2001; 2001US-00907969.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Bangur CS,  Retter MW,  Fanger GR,  Hill P;
XX
DR   WPI; 2003-897152/82.
XX
PT   Oncogenic nucleic acids useful for the prevention, diagnosis and
PT   treatment of breast cancer.
XX
PS   Example 1; SEQ ID NO 99; 399pp; English.
XX
CC   The invention describes nucleic acids (I) and the polypeptides (II) they
CC   encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC   treating diseases related to their aberrant expression i.e. breast
CC   cancers. For example, (I) and (II) may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of (II) by expressing
CC   inactive proteins or to supplement the patients own production of (II).
CC   Additionally, (I) may be used to produce (II), by inserting (I) into a
CC   host cell and culturing the cell to express the protein (II). (I) And its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids in
CC   samples, and therefore which patients may be in need of restorative
CC   therapy. The host cell may also be used as antigens in the production of
CC   antibodies against (II) and in assays to identify modulators of (II)'s
CC   expression and activity. The anti-(II) antibodies, agonists and
CC   antagonists may be used to regulate expression and activity and as
CC   diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC   immunoassay). This sequence represents a polynucleotide encoding a
CC   secreted ovarian carcinoma antigen.
XX
SQ   Sequence 636 BP; 137 A; 158 C; 158 G; 165 T; 0 U; 18 Other;

  Query Match             5.6%;  Score 123;  DB 10;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-50;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
              |||
Db        459 GAA 457


RESULT 192
ADG46343/c
ID   ADG46343 standard; cDNA; 636 BP.
XX
AC   ADG46343;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human ovarian carcinoma polynucleotide #99.
XX
KW   Human; ovarian carcinoma; gene; ss; O8E; ovarian cancer;
KW   secreted tumour antigen; cytostatic; O772P.
XX
OS   Homo sapiens.
XX
PN   US2003165504-A1.
XX
PD   04-SEP-2003.
XX
PF   04-APR-2001; 2001US-00827271.
XX
PR   17-DEC-1998;   98US-00215681.
PR   17-DEC-1998;   98US-00216003.
PR   23-JUN-1999;   99US-00338933.
PR   24-SEP-1999;   99US-00404879.
PR   17-JUL-2000; 2000US-00617747.
PR   10-AUG-2000; 2000US-00636801.
PR   20-SEP-2000; 2000US-00667857.
XX
PA   (RETT/) RETTER M W.
PA   (FANG/) FANGER G R.
XX
PI   Retter MW,  Fanger GR;
XX
DR   WPI; 2003-898035/82.
XX
PT   New isolated O8E or O772P polypeptides, useful for diagnosing,
PT   preventing, treating and monitoring cancer, e.g. ovarian cancer,
PT   stimulating the immune response in patient.
XX
PS   Example 1; SEQ ID NO 99; 290pp; English.
XX
CC   The invention relates to human ovarian carcinoma polypeptides, designated
CC   O8E or O772P, and the polynucleotides encoding them. The invention also
CC   relates to methods for inhibiting the development of cancer, e.g. ovarian
CC   cancer in a patient, methods for stimulating and/or expanding T cells and
CC   methods for identifying secreted tumour antigens. The polypeptides,
CC   compositions, antibodies to the polypeptides and methods are useful for
CC   diagnosing, preventing, treating and monitoring cancer, e.g. ovarian
CC   cancer. The composition is particularly useful for stimulating an immune
CC   response in patient. This sequence represents a human ovarian carcinoma
CC   polynucleotide of the invention.
XX
SQ   Sequence 636 BP; 137 A; 158 C; 158 G; 165 T; 0 U; 18 Other;

  Query Match             5.6%;  Score 123;  DB 10;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-50;
```

```
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db         459 GAA 457
```

RESULT 193
ADX17478/c
ID ADX17478 standard; cDNA; 636 BP.
XX
AC ADX17478;
XX
DT 21-APR-2005 (first entry)
XX
DE Human ovarian carcinoma antigen cDNA #92.
XX
KW diagnosis; cytostatic; immunostimulant; gene therapy; tumor;
KW ovarian tumor; cancer; carcinoma; antigen; ss.
XX
OS Homo sapiens.
XX
PN US2005031634-A1.
XX
PD 10-FEB-2005.
XX
PF 02-JUN-2004; 2004US-00860790.
XX
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
PR 17-JUL-2001; 2001US-00907969.
PR 17-JUL-2002; 2002US-00198053.
XX
PA (CORI-) CORIXA CORP.
XX
PI Bangur CS, Retter MW, Fanger GR, Hill P;
XX
DR WPI; 2005-151645/16.
XX
PT New ovarian carcinoma polynucleotides, preferably cDNAs, useful for
PT diagnosing, preventing and treating diseases, such as ovarian cancer, and
PT for eliciting humoral and/or cellular immune response.
XX
PS Example 1; SEQ ID NO 99; 398pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising a sequence
CC of, a sequence hybridizing under highly stringent conditions to, or
CC having at least 75, or 90% identity to a 849 or 1399 bp, fully defined in
CC the specification, its complement or degenerate variants, or a sequence
CC of at least 20 contiguous residues of the 849 or 1399 bp sequence. The
CC polynucleotides and polypeptides are useful for diagnosing, preventing
CC and treating diseases, such as ovarian cancer, and for eliciting humoral
CC and/or cellular immune response. This sequence corresponds to an ovarian
CC carcinoma antigen cDNA of the invention.
XX
SQ Sequence 636 BP; 137 A; 158 C; 158 G; 165 T; 0 U; 18 Other;

```
  Query Match             5.6%;  Score 123;  DB 14;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-50;

  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db         459 GAA 457
```

RESULT 194
AAI69048/c
ID AAI69048 standard; DNA; 184 BP.
XX
AC AAI69048;
XX
DT 29-JAN-2002 (first entry)
XX
DE Activated T-cell derived DNA fragment #184.
XX
KW Activated T-cell; immunosuppressive; immunostimulant; antiinflammatory;
KW cytostatic; gene therapy; vaccine; allergen; transplant rejection;
KW guest versus host disease; malignant disease; ds.
XX
OS Homo sapiens.
XX
PN DE10021834-A1.
XX
PD 15-NOV-2001.
XX
PF 06-MAY-2000; 2000DE-01021834.
XX
PR 06-MAY-2000; 2000DE-01021834.
XX
PA (LYNX-) LYNX THERAPEUTICS GMBH.
XX
PI Kramer MD, Winter H, Reinartz J;
XX
DR WPI; 2002-027320/04.
XX
PT New mRNA indicative of T cell activation and functional status, useful
PT for diagnosis and therapy e.g. of autoimmunity or transplant rejection.
XX
PS Claim 1; Page 52; 94pp; German.
XX
CC This sequence represents a novel messenger RNA, (mRNA), (I), for use as
CC indicator of the activation and functional status of T cells, that have
CC increased or reduced expression, and are present at higher or lower
CC concentration, in activated T cells, relative to normal or resting cells,
CC where (I) hybridizes to any of 334 sequences, reproduced, or their
CC derivatives, complements or fragments. The products of the invention have
CC immunosuppressive, immunostimulant, antiinflammatory and cytostatic
CC activity and can be used for gene therapy. The polynucleotides of the
CC invention are used: (i) as reagent for detecting activation/functional
CC status of T cells, for diagnosis, therapy, modulation or control of the
CC status, in cases of (auto)immunity (against microorganisms, vaccines or
CC allergens); transplant rejection; immunologically-related inflammation;
CC immunosuppression; immune deficiency; guest versus host disease, and
CC malignant diseases of the immune system; (ii) for identifying agents,
CC potential pharmaceuticals, that bind to (II) or derived polypeptides
CC (III); (iii) to prepare kits for measuring gene expression profiles in
CC isolated immune, especially T, cells; (iv) to raise antibodies (Ab)
CC directed against (III); and (v) to prepare binding molecules (IV)
CC specific for (II). Ab and (IV) are also useful for detecting and
CC modulating the activation and functional status of T cells. AAI68865-
CC AAI69198 represent the activated T-cell derived polynucleotide fragments
CC described in the method of the invention
XX

```
SQ   Sequence 184 BP; 53 A; 38 C; 42 G; 51 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 121;  DB 6;  Length 184;
  Best Local Similarity   99.4%;  Pred. No. 1.7e-49;
  Matches  171;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2021 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 2080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 125

Qy       2081 AGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAA 2140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTTTAAGTTTCCCCTTTTAA 65

Qy       2141 GGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 13


RESULT 195
ACH41117
ID   ACH41117 standard; cDNA; 469 BP.
XX
AC   ACH41117;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human foetal brain cDNA #2484.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 28329; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 469 BP; 165 A; 103 C; 96 G; 97 T; 0 U; 8 Other;

  Query Match             5.4%;  Score 118;  DB 9;  Length 469;
  Best Local Similarity   99.4%;  Pred. No. 5.1e-48;
  Matches  168;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2024 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 2083
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 245

Qy       2084 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 2143
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         246 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAATTTCCCCTTTTAAGGT 305

Qy       2144 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 354


RESULT 196
AAV60065/c
ID   AAV60065 standard; cDNA to mRNA; 795 BP.
XX
AC   AAV60065;
XX
DT   25-NOV-1998  (first entry)
XX
DE   Nucleic acid E67 homologous to human ubiquitin gene.
XX
KW   Human homologue; yeast RAD50; Drosophila Septin-2; Acyl-CoA synthetase;
KW   immunomodulatory activity; identification; activated T-cell; ds.
XX
OS   Homo sapiens.
XX
PN   WO9838306-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1997;   97WO-US003159.
XX
PR   27-FEB-1997;   97WO-US003159.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Dolganov G;
XX
DR   WPI; 1998-481207/41.
XX
PT   Novel human immunomodulatory poly:peptide(s) - have homology to the yeast
PT   RAD50 or Drosophila Septin-2 proteins.
XX
PS   Disclosure; Page 127; 155pp; English.
XX
CC   The present sequence represents a nucleic acid sequence identified in the
CC   course of the invention, and 90% homologous to a human ubiquitin gene.
CC   The specification also describes sequences encoding human homologues of
CC   the yeast RAD50, the Drosophila Septin-2 and Acyl-CoA synthetase. The
CC   proteins have immunomodulatory activity. The nucleic acids and proteins
CC   can be used to identify activated T-cells in a sample population. They
CC   can also be used to isolate and identify sequences encoding other
CC   proteins or other compounds having immunomodulatory activity
XX
SQ   Sequence 795 BP; 152 A; 209 C; 211 G; 214 T; 0 U; 9 Other;

  Query Match             4.7%;  Score 104;  DB 2;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-41;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
                 |||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6
```

RESULT 197
ABQ59388
ID ABQ59388 standard; cDNA; 520 BP.
XX
AC ABQ59388;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:3083.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 520 BP; 137 A; 133 C; 111 G; 113 T; 0 U; 26 Other;

Query Match 4.6%; Score 100; DB 6; Length 520;
Best Local Similarity 100.0%; Pred. No. 4.4e-39;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 91

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              |||||||||||||||||||||||||||||||||||||||||
Db         92 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 131
```

Search completed: March 6, 2006, 20:51:18
Job time : 1241 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 20:57:03 ; Search time 1792 Seconds
(without alignments)
10115.221 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 100

Total number of hits satisfying chosen parameters: 206

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2192	100.0	2192	6	US-10-170-385-378	Sequence 378, App
2	2192	100.0	2192	8	US-10-684-422-1	Sequence 1, Appli
3	2177	99.3	2201	8	US-10-769-476-3	Sequence 3, Appli
4	2177	99.3	2210	8	US-10-769-476-2	Sequence 2, Appli
5	2128	97.1	3055	9	US-10-956-157-4989	Sequence 4989, Ap
6	1348	61.5	1400	9	US-10-956-157-10224	Sequence 10224, A
7	1283	58.5	2309	3	US-09-969-708-243	Sequence 243, App
8	1283	58.5	2309	9	US-10-843-641A-7714	Sequence 7714, Ap
9	987	45.0	1055	9	US-10-956-157-4988	Sequence 4988, Ap
10	987	45.0	1055	9	US-10-956-157-10223	Sequence 10223, A
11	719	32.8	1300	6	US-10-094-749-1220	Sequence 1220, Ap
12	476	21.7	693	6	US-10-264-049-1843	Sequence 1843, Ap
13	449	20.5	547	3	US-09-738-973-171	Sequence 171, App
14	449	20.5	547	3	US-09-854-133-171	Sequence 171, App
15	449	20.5	547	5	US-10-144-649A-171	Sequence 171, App
16	433	19.8	460	3	US-09-922-217-163	Sequence 163, App
17	433	19.8	460	3	US-09-833-263-163	Sequence 163, App
18	433	19.8	460	5	US-10-025-380-163	Sequence 163, App
19	430	19.6	494	3	US-09-738-973-275	Sequence 275, App
20	430	19.6	494	3	US-09-854-133-275	Sequence 275, App
21	430	19.6	494	5	US-10-144-649A-275	Sequence 275, App
22	426	19.4	483	3	US-09-738-973-141	Sequence 141, App
23	426	19.4	483	3	US-09-854-133-141	Sequence 141, App
24	426	19.4	483	5	US-10-144-649A-141	Sequence 141, App
25	411	18.8	412	3	US-09-854-133-692	Sequence 692, App
26	411	18.8	412	5	US-10-144-649A-692	Sequence 692, App
27	410	18.7	434	3	US-09-738-973-248	Sequence 248, App
28	410	18.7	434	3	US-09-854-133-248	Sequence 248, App
29	410	18.7	434	5	US-10-144-649A-248	Sequence 248, App
30	408	18.6	442	3	US-09-918-995-23672	Sequence 23672, A
31	402	18.3	470	7	US-10-242-535A-41646	Sequence 41646, A
32	402	18.3	470	7	US-10-085-783A-41646	Sequence 41646, A
33	398	18.2	1187	3	US-09-925-298-43	Sequence 43, Appl
34	398	18.2	1187	5	US-10-102-806-43	Sequence 43, Appl
35	376	17.2	389	3	US-09-918-995-6326	Sequence 6326, Ap
36	376	17.2	506	3	US-09-918-995-20662	Sequence 20662, A
37	370	16.9	518	3	US-09-884-441-152	Sequence 152, App
38	370	16.9	518	3	US-09-907-969-152	Sequence 152, App
39	370	16.9	518	3	US-09-827-271-152	Sequence 152, App
40	370	16.9	518	6	US-10-198-053-152	Sequence 152, App
41	370	16.9	518	8	US-10-860-790-152	Sequence 152, App
42	327	14.9	400	7	US-10-242-535A-12447	Sequence 12447, A
43	327	14.9	400	7	US-10-085-783A-12447	Sequence 12447, A
44	327	14.9	401	7	US-10-242-535A-33895	Sequence 33895, A
45	327	14.9	401	7	US-10-085-783A-33895	Sequence 33895, A
46	323	14.7	472	7	US-10-242-535A-42047	Sequence 42047, A
47	323	14.7	472	7	US-10-085-783A-42047	Sequence 42047, A
48	304	13.9	389	7	US-10-242-535A-35539	Sequence 35539, A
49	304	13.9	389	7	US-10-085-783A-35539	Sequence 35539, A
50	303	13.8	308	7	US-10-242-535A-29244	Sequence 29244, A
51	303	13.8	308	7	US-10-085-783A-29244	Sequence 29244, A
52	303	13.8	412	3	US-09-922-217-660	Sequence 660, App
53	303	13.8	412	3	US-09-833-263-660	Sequence 660, App
54	303	13.8	412	5	US-10-025-380-660	Sequence 660, App
55	300	13.7	482	3	US-09-918-995-9561	Sequence 9561, Ap
56	299	13.6	312	7	US-10-242-535A-34175	Sequence 34175, A
57	299	13.6	312	7	US-10-085-783A-34175	Sequence 34175, A
58	299	13.6	400	7	US-10-242-535A-21865	Sequence 21865, A
59	299	13.6	400	7	US-10-085-783A-21865	Sequence 21865, A
60	294	13.4	393	3	US-09-918-995-3620	Sequence 3620, Ap
c 61	288	13.1	465	5	US-10-027-632-3839	Sequence 3839, Ap
c 62	288	13.1	465	6	US-10-027-632-3839	Sequence 3839, Ap
63	281	12.8	332	7	US-10-242-535A-49271	Sequence 49271, A
64	281	12.8	332	7	US-10-085-783A-49271	Sequence 49271, A
65	276	12.6	361	7	US-10-242-535A-31879	Sequence 31879, A
66	276	12.6	361	7	US-10-085-783A-31879	Sequence 31879, A
67	275	12.5	399	3	US-09-918-995-7727	Sequence 7727, Ap
68	271	12.4	396	7	US-10-242-535A-33670	Sequence 33670, A
69	271	12.4	396	7	US-10-085-783A-33670	Sequence 33670, A
70	270	12.3	284	5	US-10-004-427-52	Sequence 52, Appl
71	270	12.3	284	9	US-10-984-752-52	Sequence 52, Appl
72	270	12.3	288	7	US-10-242-535A-12261	Sequence 12261, A
73	270	12.3	288	7	US-10-085-783A-12261	Sequence 12261, A
74	260	11.9	411	3	US-09-918-995-8065	Sequence 8065, Ap
75	259	11.8	345	7	US-10-242-535A-43141	Sequence 43141, A
76	259	11.8	345	7	US-10-085-783A-43141	Sequence 43141, A
77	258	11.8	339	7	US-10-242-535A-31872	Sequence 31872, A
78	258	11.8	339	7	US-10-085-783A-31872	Sequence 31872, A
79	258	11.8	345	7	US-10-242-535A-50008	Sequence 50008, A
80	258	11.8	345	7	US-10-085-783A-50008	Sequence 50008, A
81	255	11.6	337	7	US-10-242-535A-30623	Sequence 30623, A
82	255	11.6	337	7	US-10-085-783A-30623	Sequence 30623, A
83	248	11.3	585	3	US-09-884-441-45	Sequence 45, Appl
84	248	11.3	585	3	US-09-907-969-45	Sequence 45, Appl
85	248	11.3	585	3	US-09-827-271-45	Sequence 45, Appl
86	248	11.3	585	6	US-10-198-053-45	Sequence 45, Appl
87	248	11.3	585	8	US-10-860-790-45	Sequence 45, Appl
c 88	244	11.1	467	8	US-10-357-930-45895	Sequence 45895, A
89	242	11.0	329	7	US-10-242-535A-10811	Sequence 10811, A
90	242	11.0	329	7	US-10-085-783A-10811	Sequence 10811, A
91	238	10.9	337	7	US-10-242-535A-51059	Sequence 51059, A
92	238	10.9	337	7	US-10-085-783A-51059	Sequence 51059, A
93	236	10.8	474	7	US-10-242-535A-28641	Sequence 28641, A
94	236	10.8	474	7	US-10-085-783A-28641	Sequence 28641, A
95	235	10.7	286	7	US-10-242-535A-33237	Sequence 33237, A
96	235	10.7	286	7	US-10-085-783A-33237	Sequence 33237, A

	97	233	10.6	237	7	US-10-242-535A-27365 Sequence 27365, A
	98	233	10.6	237	7	US-10-085-783A-27365 Sequence 27365, A
	99	232	10.6	232	7	US-10-242-535A-34542 Sequence 34542, A
	100	232	10.6	232	7	US-10-085-783A-34542 Sequence 34542, A
c	101	228	10.4	228	3	US-09-796-692-3431 Sequence 3431, Ap
c	102	228	10.4	228	5	US-10-040-862-3431 Sequence 3431, Ap
c	103	228	10.4	228	6	US-10-057-475B-3431 Sequence 3431, Ap
c	104	228	10.4	228	6	US-10-154-884B-3431 Sequence 3431, Ap
c	105	228	10.4	228	8	US-10-764-324-3431 Sequence 3431, Ap
	106	223	10.2	232	7	US-10-242-535A-49909 Sequence 49909, A
	107	223	10.2	232	7	US-10-085-783A-49909 Sequence 49909, A
	108	218	9.9	403	7	US-10-242-535A-42741 Sequence 42741, A
	109	218	9.9	403	7	US-10-085-783A-42741 Sequence 42741, A
	110	217	9.9	299	8	US-10-425-115-23758 Sequence 23758, A
	111	217	9.9	372	7	US-10-242-535A-25883 Sequence 25883, A
	112	217	9.9	372	7	US-10-085-783A-25883 Sequence 25883, A
	113	213	9.7	249	7	US-10-242-535A-27257 Sequence 27257, A
	114	213	9.7	249	7	US-10-085-783A-27257 Sequence 27257, A
	115	211	9.6	440	3	US-09-876-889-277 Sequence 277, App
	116	210	9.6	261	7	US-10-242-535A-30987 Sequence 30987, A
	117	210	9.6	261	7	US-10-085-783A-30987 Sequence 30987, A
	118	210	9.6	330	7	US-10-242-535A-8262 Sequence 8262, Ap
	119	210	9.6	330	7	US-10-085-783A-8262 Sequence 8262, Ap
	120	207	9.4	228	3	US-09-938-864-384 Sequence 384, App
	121	207	9.4	228	5	US-10-125-635A-384 Sequence 384, App
	122	207	9.4	228	5	US-10-121-988-80 Sequence 80, Appl
	123	207	9.4	228	5	US-10-002-603-384 Sequence 384, App
	124	207	9.4	228	6	US-10-200-562-80 Sequence 80, Appl
	125	207	9.4	228	6	US-10-237-551-80 Sequence 80, Appl
	126	207	9.4	228	6	US-10-195-835-384 Sequence 384, App
	127	207	9.4	228	6	US-10-286-333-384 Sequence 384, App
	128	207	9.4	228	6	US-10-244-830-384 Sequence 384, App
	129	207	9.4	228	7	US-10-427-717-384 Sequence 384, App
	130	207	9.4	228	7	US-10-648-780-384 Sequence 384, App
	131	207	9.4	228	9	US-10-945-050-80 Sequence 80, Appl
	132	207	9.4	231	8	US-10-475-203A-1 Sequence 1, Appli
c	133	207	9.4	609	3	US-09-871-161-169 Sequence 169, App
	134	205	9.4	332	3	US-09-925-299-269 Sequence 269, App
	135	205	9.4	332	3	US-09-925-299-269 Sequence 269, App
	136	204	9.3	384	7	US-10-242-535A-3777 Sequence 3777, Ap
	137	204	9.3	384	7	US-10-085-783A-3777 Sequence 3777, Ap
	138	203	9.3	480	6	US-10-360-053-9 Sequence 9, Appli
	139	203	9.3	504	6	US-10-360-053-11 Sequence 11, Appl
	140	202	9.2	261	7	US-10-242-535A-36466 Sequence 36466, A
	141	202	9.2	261	7	US-10-085-783A-36466 Sequence 36466, A
	142	202	9.2	490	7	US-10-242-535A-40075 Sequence 40075, A
	143	202	9.2	490	7	US-10-085-783A-40075 Sequence 40075, A
c	144	201	9.2	201	3	US-09-796-692-3812 Sequence 3812, Ap
c	145	201	9.2	201	5	US-10-040-862-3812 Sequence 3812, Ap
c	146	201	9.2	201	6	US-10-057-475B-3812 Sequence 3812, Ap
c	147	201	9.2	201	6	US-10-154-884B-3812 Sequence 3812, Ap
c	148	201	9.2	201	8	US-10-764-324-3812 Sequence 3812, Ap
	149	201	9.2	471	3	US-09-918-995-22118 Sequence 22118, A
	150	200	9.1	216	7	US-10-242-535A-35923 Sequence 35923, A
	151	200	9.1	216	7	US-10-085-783A-35923 Sequence 35923, A
	152	200	9.1	367	7	US-10-242-535A-13820 Sequence 13820, A
	153	200	9.1	367	7	US-10-085-783A-13820 Sequence 13820, A
	154	192	8.8	294	7	US-10-242-535A-50750 Sequence 50750, A
	155	192	8.8	294	7	US-10-085-783A-50750 Sequence 50750, A
c	156	191	8.7	605	3	US-09-884-441-95 Sequence 95, Appl
c	157	191	8.7	605	3	US-09-907-969-95 Sequence 95, Appl
c	158	191	8.7	605	3	US-09-827-271-95 Sequence 95, Appl
c	159	191	8.7	605	6	US-10-198-053-95 Sequence 95, Appl
c	160	191	8.7	605	8	US-10-860-790-95 Sequence 95, Appl
	161	191	8.7	697	3	US-09-884-441-100 Sequence 100, App
	162	191	8.7	697	3	US-09-907-969-100 Sequence 100, App
	163	191	8.7	697	3	US-09-827-271-100 Sequence 100, App
	164	191	8.7	697	6	US-10-198-053-100 Sequence 100, App
	165	191	8.7	697	8	US-10-860-790-100 Sequence 100, App
	166	183	8.3	355	7	US-10-242-535A-24536 Sequence 24536, A
	167	183	8.3	355	7	US-10-085-783A-24536 Sequence 24536, A
	168	182	8.3	321	7	US-10-242-535A-5700 Sequence 5700, Ap
	169	182	8.3	321	7	US-10-085-783A-5700 Sequence 5700, Ap
	170	176	8.0	227	3	US-09-777-564-1048 Sequence 1048, Ap
	171	176	8.0	227	5	US-10-015-219-1048 Sequence 1048, Ap
c	172	171	7.8	294	8	US-10-357-930-16069 Sequence 16069, A
	173	161	7.3	273	7	US-10-242-535A-36526 Sequence 36526, A
	174	161	7.3	273	7	US-10-085-783A-36526 Sequence 36526, A
	175	160	7.3	211	7	US-10-242-535A-50471 Sequence 50471, A
	176	160	7.3	211	7	US-10-085-783A-50471 Sequence 50471, A
c	177	152	6.9	509	3	US-09-969-034-3133 Sequence 3133, Ap
	178	143	6.5	400	7	US-10-242-535A-8748 Sequence 8748, Ap
	179	143	6.5	400	7	US-10-085-783A-8748 Sequence 8748, Ap
	180	141	6.4	794	6	US-10-264-049-1008 Sequence 1008, Ap
c	181	137	6.2	240	3	US-09-777-564-1207 Sequence 1207, Ap
c	182	137	6.2	240	5	US-10-015-219-1207 Sequence 1207, Ap
	183	135	6.2	208	7	US-10-242-535A-45031 Sequence 45031, A
	184	135	6.2	208	7	US-10-085-783A-45031 Sequence 45031, A
	185	133	6.1	133	9	US-10-913-447-67 Sequence 67, Appl
	186	133	6.1	133	10	US-11-034-010-33 Sequence 33, Appl
	187	132	6.0	374	7	US-10-242-535A-34635 Sequence 34635, A
	188	132	6.0	374	7	US-10-085-783A-34635 Sequence 34635, A
	189	125	5.7	168	7	US-10-242-535A-4391 Sequence 4391, Ap
	190	125	5.7	168	7	US-10-085-783A-4391 Sequence 4391, Ap
c	191	123	5.6	636	3	US-09-884-441-99 Sequence 99, Appl
c	192	123	5.6	636	3	US-09-907-969-99 Sequence 99, Appl
c	193	123	5.6	636	3	US-09-827-271-99 Sequence 99, Appl
c	194	123	5.6	636	6	US-10-198-053-99 Sequence 99, Appl
c	195	123	5.6	636	8	US-10-860-790-99 Sequence 99, Appl
	196	120	5.5	341	7	US-10-242-535A-34587 Sequence 34587, A
	197	120	5.5	341	7	US-10-085-783A-34587 Sequence 34587, A
	198	118	5.4	469	3	US-09-918-995-28329 Sequence 28329, A
	199	112	5.1	112	7	US-10-242-535A-28125 Sequence 28125, A
	200	112	5.1	112	7	US-10-085-783A-28125 Sequence 28125, A
	201	109	5.0	167	7	US-10-242-535A-4938 Sequence 4938, Ap
	202	109	5.0	167	7	US-10-085-783A-4938 Sequence 4938, Ap
c	203	104	4.7	795	6	US-10-393-602-140 Sequence 140, App
	204	100	4.6	107	7	US-10-242-535A-13237 Sequence 13237, A
	205	100	4.6	107	7	US-10-085-783A-13237 Sequence 13237, A
	206	100	4.6	520	3	US-09-969-034-3083 Sequence 3083, Ap

ALIGNMENTS

RESULT 1
US-10-170-385-378
; Sequence 378, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 378
; LENGTH: 2192
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-378

Query Match 100.0%; Score 2192; DB 6; Length 2192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680

Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
```

```
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192


RESULT 2
US-10-684-422-1
; Sequence 1, Application US/10684422
; Publication No. US20040229233A1
; GENERAL INFORMATION:
;  APPLICANT: ABURATANI, Hiroyuki
;  APPLICANT:  YAMAMOTO, Shogo
;  TITLE OF INVENTION: Human housekeeping genes and human tissue-specific genes
;  FILE REFERENCE: 113991
;  CURRENT APPLICATION NUMBER: US/10/684,422
;  CURRENT FILING DATE:  2002-10-15
;  PRIOR APPLICATION NUMBER: US 60/418,614
;  PRIOR FILING DATE: 2002-10-16
;  NUMBER OF SEQ ID NOS: 332
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 2192
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (69)..(2126)
US-10-684-422-1

  Query Match            100.0%;  Score 2192;  DB 8;  Length 2192;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2192;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720

Qy        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780

Qy        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840

Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCA 1200

Qy       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGG 1260

Qy       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1320

Qy       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1380

Qy       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440

Qy       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1500

Qy       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1560

Qy       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 1620

Qy       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGA 1680
```

```
Qy       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1740

Qy       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1800

Qy       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1860

Qy       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1920

Qy       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 1980

Qy       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 2040

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||
Db       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
```

RESULT 3
US-10-769-476-3
; Sequence 3, Application US/10769476
; Publication No. US20050014168A1
; GENERAL INFORMATION:
; APPLICANT: Arcturus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; APPLICANT: Ma, Xiao-Jun
; APPLICANT: Baer, Thomas M.
; TITLE OF INVENTION: 3' Biased Microarrays
; FILE REFERENCE: 022041-001210US
; CURRENT APPLICATION NUMBER: US/10/769,476
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/475,812
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-769-476-3

Query Match 99.3%; Score 2177; DB 8; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
                |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 4
US-10-769-476-2
; Sequence 2, Application US/10769476
; Publication No. US20050014168A1
; GENERAL INFORMATION:
; APPLICANT: Arcturus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; APPLICANT: Ma, Xiao-Jun
; APPLICANT: Baer, Thomas M.
; TITLE OF INVENTION: 3' Biased Microarrays
; FILE REFERENCE: 022041-001210US
; CURRENT APPLICATION NUMBER: US/10/769,476
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/475,812
; PRIOR FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2210
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-769-476-2

Query Match 99.3%; Score 2177; DB 8; Length 2210;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 69

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 129

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 189

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 249

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 309

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 369

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 429

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 489

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 549

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 609

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 669

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 729

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 789
```

```
Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 849

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 909

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 969

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1029

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1089

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1149

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1209

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1269

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1329

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1389

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1449

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1509

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1569

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1629

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1630 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1689

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1690 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1749

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1750 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1809

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1810 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1869

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1870 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1929

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1930 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1989

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1990 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2049

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2050 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2109

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2110 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2169

Qy       2176 AAAGTTGTTGCATTCCC 2192
              |||||||||||||||||
Db       2170 AAAGTTGTTGCATTCCC 2186
```

RESULT 5
US-10-956-157-4989
; Sequence 4989, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WI
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4989
; LENGTH: 3055
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-4989

Query Match 97.1%; Score 2128; DB 9; Length 3055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         876 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 935

Qy         125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         936 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 995

Qy         185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         996 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 1055

Qy         245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1056 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 1115

Qy         305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1116 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 1175

Qy         365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1176 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 1235
```

```
Qy       425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
Db      1236 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 1295

Qy       485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
Db      1296 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1355

Qy       545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
Db      1356 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1415

Qy       605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
Db      1416 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 1475

Qy       665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
Db      1476 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 1535

Qy       725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
Db      1536 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 1595

Qy       785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
Db      1596 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 1655

Qy       845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
Db      1656 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 1715

Qy       905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
Db      1716 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 1775

Qy       965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
Db      1776 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1835

Qy      1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
Db      1836 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1895

Qy      1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
Db      1896 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1955

Qy      1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
Db      1956 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 2015

Qy      1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
Db      2016 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 2075

Qy      1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
Db      2076 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 2135

Qy      1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
Db      2136 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 2195

Qy      1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
Db      2196 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 2255

Qy      1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
Db      2256 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 2315

Qy      1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
Db      2316 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 2375

Qy      1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
Db      2376 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 2435

Qy      1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
Db      2436 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 2495

Qy      1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
Db      2496 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 2555

Qy      1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
Db      2556 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 2615

Qy      1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
Db      2616 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 2675

Qy      1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1924
Db      2676 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 2735

Qy      1925 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1984
Db      2736 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 2795

Qy      1985 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2044
Db      2796 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2855

Qy      2045 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2104
Db      2856 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2915

Qy      2105 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2164
Db      2916 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2975

Qy      2165 TTCCTTTCAATAAAGTTGTTGCATTCCC 2192
Db      2976 TTCCTTTCAATAAAGTTGTTGCATTCCC 3003
```

RESULT 6
US-10-956-157-10224
; Sequence 10224, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10224
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-10224

Query Match 61.5%; Score 1348; DB 9; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
```

```
Db          1 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 60

Qy        905 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 120

Qy        965 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 180

Qy       1025 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 240

Qy       1085 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 300

Qy       1145 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGG 360

Qy       1205 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 420

Qy       1265 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 480

Qy       1325 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 540

Qy       1385 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 600

Qy       1445 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 660

Qy       1505 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 720

Qy       1565 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 780

Qy       1625 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 840

Qy       1685 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 900

Qy       1745 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 960

Qy       1805 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 1020

Qy       1865 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1080

Qy       1925 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 1140

Qy       1985 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 2044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 1200

Qy       2045 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 2104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCT 1260

Qy       2105 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 2164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACT 1320

Qy       2165 TTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||
Db       1321 TTCCTTTCAATAAAGTTGTTGCATTCCC 1348


RESULT 7
US-09-969-708-243
; Sequence 243, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
;  APPLICANT: Augustus, Meena
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-70
;  CURRENT APPLICATION NUMBER: US/09/969,708
;  CURRENT FILING DATE:  2001-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,606
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,608
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: US/60/237,425
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 658
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 243
;   LENGTH: 2309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-969-708-243

  Query Match             58.5%;  Score 1283;  DB 3;  Length 2309;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516

Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
```

```
Db        517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576

Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636

Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696

Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756

Qy       1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816

Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876

Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936

Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996

Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056

Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116

Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176

Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236

Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296

Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356

Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416

Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
              |||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643
```

```
RESULT 8
US-10-843-641A-7714
; Sequence 7714, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
;  APPLICANT: Avalon Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-189
;  CURRENT APPLICATION NUMBER: US/10/843,641A
;  CURRENT FILING DATE:  2004-05-12
;  PRIOR APPLICATION NUMBER: US/09/873,367
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US/09/954,531
;  PRIOR FILING DATE: 2001-09-18
;  PRIOR APPLICATION NUMBER: US/09/954,456
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,436
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/962,832
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/09/964,824
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US/09/967,768
;  PRIOR FILING DATE: 2001-09-28
;  PRIOR APPLICATION NUMBER: US/09/968,007
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,347
;  PRIOR FILING DATE: 2001-10-02
;  PRIOR APPLICATION NUMBER: US/09/969,708
;  PRIOR FILING DATE: 2001-10-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 8447
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7714
;   LENGTH: 2309
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-843-641A-7714

  Query Match            58.5%;  Score 1283;  DB 9;  Length 2309;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 216

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 276

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 336

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 396

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db        397 ACCATCACCCTTGAGGTCGAGCCCAGTGACACAATCGAGAACGTCAAGGCAAAGATCCAA 456

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 516

Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        517 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 576
Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
Db        577 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 636
Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
Db        637 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 696
Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
Db        697 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 756
Qy       1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
Db        757 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 816
Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
Db        817 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 876
Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
Db        877 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 936
Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
Db        937 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 996
Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
Db        997 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1056
Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
Db       1057 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1116
Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
Db       1117 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1176
Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
Db       1177 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1236
Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
Db       1237 GAGTCCACCCTGCACCTGGTTCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1296
Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
Db       1297 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1356
Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
Db       1357 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1416
Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
Db       1417 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1476
Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
Db       1477 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1536
Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
Db       1537 ACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1596
Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
Db       1597 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 1643
```

```
RESULT 9
US-10-956-157-4988
; Sequence 4988, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WI
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4988
;   LENGTH: 1055
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-4988

  Query Match            45.0%;  Score 987;  DB 9;  Length 1055;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  987;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
Db         69 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 128
Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
Db        129 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 188
Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
Db        189 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 248
Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
Db        249 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 308
Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
Db        309 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 368
Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
Db        369 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 428
Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
Db        429 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 488
Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
Db        489 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 548
Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
Db        549 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 608
Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
Db        609 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 668
Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
Db        669 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 728
Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
Db        729 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 788
```

```
Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 848

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 2045
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 908

Qy       2046 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 2105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 968

Qy       2106 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 2165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 1028

Qy       2166 TCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||
Db       1029 TCCTTTCAATAAAGTTGTTGCATTCCC 1055
```

RESULT 10
US-10-956-157-10223
; Sequence 10223, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10223
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-10223

Query Match 45.0%; Score 987; DB 9; Length 1055;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          69 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 128

Qy       1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         129 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 188

Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         189 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 248

Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         249 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 308

Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         309 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 368

Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         369 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 428

Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         429 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 488

Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         489 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 548

Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         549 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 608

Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         609 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 668

Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         669 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 728

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         729 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 788

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         789 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 848

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 2045
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         849 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 908

Qy       2046 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 2105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         909 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 968

Qy       2106 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 2165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         969 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 1028

Qy       2166 TCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||
Db       1029 TCCTTTCAATAAAGTTGTTGCATTCCC 1055
```

RESULT 11
US-10-094-749-1220
; Sequence 1220, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14

```
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1220
;   LENGTH: 1300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-094-749-1220

  Query Match            32.8%;  Score 719;  DB 6;  Length 1300;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  769;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 316

Qy        408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 376

Qy        468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 436

Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 496

Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 556

Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 616

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 676

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 736

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        737 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 796

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 856

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 916

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        917 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 976

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1117
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        977 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGC 1026


RESULT 12
US-10-264-049-1843
; Sequence 1843, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1843
;   LENGTH: 693
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (39)..(39)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (522)..(522)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (580)..(580)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (609)..(609)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (648)..(648)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (664)..(664)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (671)..(671)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (674)..(674)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-1843

  Query Match            21.7%;  Score 476;  DB 6;  Length 693;
  Best Local Similarity  100.0%;  Pred. No. 9.5e-246;
  Matches  476;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1257 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 105

Qy       1317 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 165

Qy       1377 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 225
```

```
Qy       1437 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 285

Qy       1497 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1556
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 345

Qy       1557 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 405

Qy       1617 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 465

Qy       1677 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 521
```

RESULT 13
US-09-738-973-171
; Sequence 171, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-171

Query Match 20.5%; Score 449; DB 3; Length 547;
Best Local Similarity 99.8%; Pred. No. 3.5e-231;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548
              ||||||||||||||||||||
Db        528 AGATCTTCGTGAAGACCCTG 547
```

RESULT 14
US-09-854-133-171
; Sequence 171, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-171

Query Match 20.5%; Score 449; DB 3; Length 547;
Best Local Similarity 99.8%; Pred. No. 3.5e-231;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548
              |||||||||||||||||||||
Db        528 AGATCTTCGTGAAGACCCTG 547


RESULT 15
US-10-144-649A-171
; Sequence 171, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C11
;  CURRENT APPLICATION NUMBER: US/10/144,649A
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 749
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 171
;   LENGTH: 547
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-144-649A-171

  Query Match             20.5%;  Score 449;  DB 5;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 3.5e-231;
  Matches  499;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548
              |||||||||||||||||||||
Db        528 AGATCTTCGTGAAGACCCTG 547


RESULT 16
US-09-922-217-163
; Sequence 163, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Smith, Carole Lynn
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Clapper, Jonathan D.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;  TITLE OF INVENTION:  OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C13
;  CURRENT APPLICATION NUMBER: US/09/922,217
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-922-217-163

  Query Match             19.8%;  Score 433;  DB 3;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-222;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 17
US-09-833-263-163
; Sequence 163, Application US/09833263
; Patent No. US20020110547A1
```

; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 163
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-833-263-163

Query Match 19.8%; Score 433; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.5e-222;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460
```

RESULT 18
US-10-025-380-163
; Sequence 163, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-025-380-163

Query Match 19.8%; Score 433; DB 5; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.5e-222;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460
```

RESULT 19
US-09-738-973-275
; Sequence 275, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9

```
;  CURRENT APPLICATION NUMBER: US/09/738,973
;  CURRENT FILING DATE:  2000-12-14
;  NUMBER OF SEQ ID NOS:  587
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 275
;   LENGTH: 494
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (379)
;   OTHER INFORMATION: n=A,T,C or G
US-09-738-973-275

  Query Match            19.6%;  Score 430;  DB 3;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 6.1e-221;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494


RESULT 20
US-09-854-133-275
; Sequence 275, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 275
;   LENGTH: 494
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (379)
;   OTHER INFORMATION: n=A,T,C or G
US-09-854-133-275

  Query Match            19.6%;  Score 430;  DB 3;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 6.1e-221;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494


RESULT 21
US-10-144-649A-275
; Sequence 275, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C11
;  CURRENT APPLICATION NUMBER: US/10/144,649A
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 749
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 275
;   LENGTH: 494
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: 379
;   OTHER INFORMATION: n = A,T,C or G
US-10-144-649A-275

  Query Match            19.6%;  Score 430;  DB 5;  Length 494;
  Best Local Similarity   99.8%;  Pred. No. 6.1e-221;
  Matches  480;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494


RESULT 22
US-09-738-973-141
; Sequence 141, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Elliot, Mark
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Kalos, Michael D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.475C9
;  CURRENT APPLICATION NUMBER: US/09/738,973
;  CURRENT FILING DATE:  2000-12-14
;  NUMBER OF SEQ ID NOS:  587
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 141
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-738-973-141

  Query Match            19.4%;  Score 426;  DB 3;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-219;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 23
US-09-854-133-141
; Sequence 141, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 141
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-854-133-141

  Query Match            19.4%;  Score 426;  DB 3;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-219;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177
```

```
Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483
```

RESULT 24
US-10-144-649A-141
; Sequence 141, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 141
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-144-649A-141

Query Match 19.4%; Score 426; DB 5; Length 483;
Best Local Similarity 100.0%; Pred. No. 8.8e-219;
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483
```

RESULT 25
US-09-854-133-692
; Sequence 692, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 692
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-692

Query Match 18.8%; Score 411; DB 3; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.1e-210;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         35 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 61

Qy         95 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 121

Qy        155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 181

Qy        215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 241

Qy        275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 301

Qy        335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 361

Qy        395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 412
```

RESULT 26
US-10-144-649A-692
; Sequence 692, Application US/10144649A
; Publication No. US20030118599A1

```
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C11
;  CURRENT APPLICATION NUMBER: US/10/144,649A
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 749
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 692
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-144-649A-692

  Query Match             18.8%;  Score 411;  DB 5;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-210;
  Matches  411;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         35 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 61

Qy         95 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 121

Qy        155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 181

Qy        215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 241

Qy        275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 301

Qy        335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 361

Qy        395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 412


RESULT 27
US-09-738-973-248
; Sequence 248, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Pling, Steven P.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Indirias, Carol Yoseph
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Elliot, Mark
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Kalos, Michael D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.475C9
;  CURRENT APPLICATION NUMBER: US/09/738,973
;  CURRENT FILING DATE:  2000-12-14
;  NUMBER OF SEQ ID NOS:  587
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 248
;   LENGTH: 434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-738-973-248

  Query Match             18.7%;  Score 410;  DB 3;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-210;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410


RESULT 28
US-09-854-133-248
; Sequence 248, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 248
;   LENGTH: 434
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-854-133-248

  Query Match             18.7%;  Score 410;  DB 3;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-210;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60
```

```
Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 29
US-10-144-649A-248
; Sequence 248, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 248
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-144-649A-248

Query Match 18.7%; Score 410; DB 5; Length 434;
Best Local Similarity 100.0%; Pred. No. 3.8e-210;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 30
US-09-918-995-23672
; Sequence 23672, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23672
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23672

Query Match 18.6%; Score 408; DB 3; Length 442;
Best Local Similarity 100.0%; Pred. No. 4.5e-209;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1722 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 94

Qy       1782 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 154

Qy       1842 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 214

Qy       1902 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 1961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 274

Qy       1962 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 2021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 334

Qy       2022 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 2081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 394

Qy       2082 GAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 2129
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 442
```

RESULT 31
US-10-242-535A-41646
; Sequence 41646, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Qsteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12

```
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41646
;   LENGTH: 470
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (425)..(425)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (464)..(464)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-41646

  Query Match             18.3%;  Score 402;  DB 7;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 7.8e-206;
  Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1717 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 60

Qy       1777 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 120

Qy       1837 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 180

Qy       1897 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 1956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 240

Qy       1957 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGA 2016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGA 300

Qy       2017 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 2076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 360

Qy       2077 AGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGG 2118
             |||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGG 402


RESULT 32
US-10-085-783A-41646
; Sequence 41646, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41646
;   LENGTH: 470
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (425)..(425)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (464)..(464)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-41646

  Query Match             18.3%;  Score 402;  DB 7;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 7.8e-206;
  Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1717 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 60

Qy       1777 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 120

Qy       1837 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 180

Qy       1897 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 1956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 240

Qy       1957 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGA 2016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGA 300

Qy       2017 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 2076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCC 360

Qy       2077 AGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGG 2118
             |||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGG 402


RESULT 33
US-09-925-298-43
; Sequence 43, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103
;  CURRENT APPLICATION NUMBER: US/09/925,298
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 1187
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (33)
```

```
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (41)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1149)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1156)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1160)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-298-43

  Query Match            18.2%;  Score 398;  DB 3;  Length 1187;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-203;
  Matches  398;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 654

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 714

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1664
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 774

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 834

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 894

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 954

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1882
             |||||||||||||||||||||||||||||||||||||||
Db        955 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 992


RESULT 34
US-10-102-806-43
; Sequence 43, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 1187
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (33)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (41)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1149)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1156)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1160)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-43

  Query Match            18.2%;  Score 398;  DB 5;  Length 1187;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-203;
  Matches  398;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 654

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 714

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1664
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 774

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 834

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 894

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 954

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1882
             |||||||||||||||||||||||||||||||||||||||
Db        955 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 992


RESULT 35
US-09-918-995-6326
; Sequence 6326, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6326
;   LENGTH: 389
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-6326

  Query Match            17.2%;  Score 376;  DB 3;  Length 389;
  Best Local Similarity  100.0%;  Pred. No. 8.2e-192;
  Matches  376;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1099 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACA 1158
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         13 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACA 72

Qy       1159 ACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 1218
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 ACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 132

Qy       1219 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTG 1278
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTG 192

Qy       1279 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGA 1338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGA 252

Qy       1339 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 1398
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 312

Qy       1399 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCC 1458
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCC 372

Qy       1459 TGACTGGTAAGACCAT 1474
             ||||||||||||||||
Db        373 TGACTGGTAAGACCAT 388
```

RESULT 36
US-09-918-995-20662
; Sequence 20662, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 20662
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(506)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-20662

Query Match 17.2%; Score 376; DB 3; Length 506;
Best Local Similarity 100.0%; Pred. No. 8.2e-192;
Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1796 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1855
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         114 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 173

Qy       1856 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1915
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         174 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 233

Qy       1916 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 1975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         234 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 293

Qy       1976 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 2035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         294 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 353

Qy       2036 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 2095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         354 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 413

Qy       2096 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 2155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         414 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 473

Qy       2156 CATTGCACTTTCCTTT 2171
             ||||||||||||||||
Db         474 CATTGCACTTTCCTTT 489
```

RESULT 37
US-09-884-441-152
; Sequence 152, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 152
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-152

Query Match 16.9%; Score 370; DB 3; Length 518;
Best Local Similarity 99.6%; Pred. No. 1.4e-188;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 38
US-09-907-969-152

```
; Sequence 152, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Hill, Paul
;  APPLICANT:  Albone, Earl
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C8
;  CURRENT APPLICATION NUMBER: US/09/907,969
;  CURRENT FILING DATE:  2001-07-17
;  NUMBER OF SEQ ID NOS: 596
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-907-969-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-188;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 39
US-09-827-271-152
; Sequence 152, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-188;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 40
US-10-198-053-152
; Sequence 152, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-152

  Query Match             16.9%;  Score 370;  DB 6;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-188;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 41
US-10-860-790-152
; Sequence 152, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-860-790-152

Query Match 16.9%; Score 370; DB 8; Length 518;
Best Local Similarity 99.6%; Pred. No. 1.4e-188;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 42
US-10-242-535A-12447
; Sequence 12447, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12447
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (66)..(66)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-12447

Query Match 14.9%; Score 327; DB 7; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.2e-165;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1856 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1915
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 133

Qy       1916 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 1975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 193

Qy       1976 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 2035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 253

Qy       2036 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 2095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 313

Qy       2096 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 2155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        314 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 373

Qy       2156 CATTGCACTTTCCTTTCAATAAAGTTG 2182
             ||||||||||||||||||||||||||||
Db        374 CATTGCACTTTCCTTTCAATAAAGTTG 400


RESULT 43
US-10-085-783A-12447
; Sequence 12447, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12447
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (66)..(66)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-12447

  Query Match             14.9%;  Score 327;  DB 7;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-165;
  Matches  327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1856 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1915
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 133

Qy       1916 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 1975
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 193

Qy       1976 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 2035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 253

Qy       2036 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 2095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCA 313

Qy       2096 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 2155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 CTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTT 373

Qy       2156 CATTGCACTTTCCTTTCAATAAAGTTG 2182
             ||||||||||||||||||||||||||||
Db        374 CATTGCACTTTCCTTTCAATAAAGTTG 400


RESULT 44
US-10-242-535A-33895
; Sequence 33895, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33895
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-33895

  Query Match             14.9%;  Score 327;  DB 7;  Length 401;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-165;
  Matches  327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1259 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 60

Qy       1319 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 120

Qy       1379 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 1438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 180

Qy       1439 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 1498
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 240

Qy       1499 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 1558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 300

Qy       1559 GAGGTTGATCTTTGCTGGGAAACAGCT 1585
             |||||||||||||||||||||||||||
Db        301 GAGGTTGATCTTTGCTGGGAAACAGCT 327


RESULT 45
US-10-085-783A-33895
; Sequence 33895, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33895
;   LENGTH: 401
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-33895
```

```
  Query Match             14.9%;  Score 327;  DB 7;  Length 401;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-165;
  Matches  327;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1259 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1318
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 60

Qy       1319 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1378
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 120

Qy       1379 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 1438
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 180

Qy       1439 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 1498
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 240

Qy       1499 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 1558
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 300

Qy       1559 GAGGTTGATCTTTGCTGGGAAACAGCT 1585
             |||||||||||||||||||||||||||
Db          301 GAGGTTGATCTTTGCTGGGAAACAGCT 327


RESULT 46
US-10-242-535A-42047
; Sequence 42047, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42047
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (324)..(324)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (378)..(379)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (391)..(391)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (414)..(414)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (422)..(422)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (430)..(430)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (456)..(456)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-42047

  Query Match             14.7%;  Score 323;  DB 7;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-163;
  Matches  323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1489 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 60

Qy       1549 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 120

Qy       1609 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 180

Qy       1669 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 240

Qy       1729 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 300

Qy       1789 GGTTGATCTTTGCTGGGAAACAG 1811
             |||||||||||||||||||||||
Db           301 GGTTGATCTTTGCTGGGAAACAG 323


RESULT 47
US-10-085-783A-42047
; Sequence 42047, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42047
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (324)..(324)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (378)..(379)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
```

; LOCATION: (391)..(391)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (414)..(414)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (422)..(422)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (430)..(430)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (456)..(456)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-42047

Query Match 14.7%; Score 323; DB 7; Length 472;
Best Local Similarity 100.0%; Pred. No. 3.1e-163;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1489 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTC 60

Qy       1549 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 120

Qy       1609 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 180

Qy       1669 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 240

Qy       1729 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGA 300

Qy       1789 GGTTGATCTTTGCTGGGAAACAG 1811
             |||||||||||||||||||||||
Db        301 GGTTGATCTTTGCTGGGAAACAG 323
```

RESULT 48
US-10-242-535A-35539
; Sequence 35539, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35539
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-35539

Query Match 13.9%; Score 304; DB 7; Length 389;
Best Local Similarity 99.7%; Pred. No. 5.6e-153;
Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1765 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 91

Qy       1825 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 1884
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 151

Qy       1885 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGG 1944
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCATCCTCGAGGTGG 211

Qy       1945 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 2004
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 271

Qy       2005 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 2064
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 331

Qy       2065 ACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 ACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 386
```

RESULT 49
US-10-085-783A-35539
; Sequence 35539, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35539
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature

```
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-35539

  Query Match             13.9%;  Score 304;  DB 7;  Length 389;
  Best Local Similarity   99.7%;  Pred. No. 5.6e-153;
  Matches  354;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1765 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 91

Qy       1825 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 1884
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         92 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 151

Qy       1885 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGG 1944
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        152 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCATCCTCGAGGTGG 211

Qy       1945 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 2004
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 271

Qy       2005 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 2064
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 CTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTG 331

Qy       2065 ACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 ACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 386


RESULT 50
US-10-242-535A-29244
; Sequence 29244, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29244
;   LENGTH: 308
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-29244

  Query Match             13.8%;  Score 303;  DB 7;  Length 308;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-152;
  Matches  303;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1888 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1947
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 64

Qy       1948 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 2007
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 124

Qy       2008 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2067
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 184

Qy       2068 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2127
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 244

Qy       2128 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2187
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 304

Qy       2188 TTC 2190
             |||
Db        305 TTC 307


RESULT 51
US-10-085-783A-29244
; Sequence 29244, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29244
;   LENGTH: 308
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-29244

  Query Match             13.8%;  Score 303;  DB 7;  Length 308;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-152;
  Matches  303;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1888 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 1947
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGC 64

Qy       1948 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 2007
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 CCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTG 124

Qy       2008 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 2067
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 184

Qy       2068 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 2127
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAG 244

Qy       2128 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 2187
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 TTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCA 304

Qy       2188 TTC 2190
             |||
Db        305 TTC 307
```

```
RESULT 52
US-09-922-217-660
; Sequence 660, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Smith, Carole Lynn
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Clapper, Jonathan D.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;  TITLE OF INVENTION:  OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C13
;  CURRENT APPLICATION NUMBER: US/09/922,217
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 660
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-922-217-660

  Query Match           13.8%;  Score 303;  DB 3;  Length 412;
  Best Local Similarity   99.7%;  Pred. No. 1.9e-152;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 53
US-09-833-263-660
; Sequence 660, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Aijun
;  APPLICANT:  Clapper, Jonathan D.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Meagher, Madeleine J.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C12
;  CURRENT APPLICATION NUMBER: US/09/833,263
;  CURRENT FILING DATE:  2001-04-10
;  NUMBER OF SEQ ID NOS: 1093
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 660
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-833-263-660

  Query Match           13.8%;  Score 303;  DB 3;  Length 412;
  Best Local Similarity   99.7%;  Pred. No. 1.9e-152;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 54
US-10-025-380-660
; Sequence 660, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Clapper, Jonathan D.
;  APPLICANT:  Skeiky, Yasir A. W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;  TITLE OF INVENTION:  OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C14
;  CURRENT APPLICATION NUMBER: US/10/025,380
;  CURRENT FILING DATE:  2001-12-19
;  NUMBER OF SEQ ID NOS: 1129
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 660
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-025-380-660

  Query Match           13.8%;  Score 303;  DB 5;  Length 412;
```

```
  Best Local Similarity   99.7%;  Pred. No. 1.9e-152;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy         126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy         186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy         246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy         306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
               ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy         366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 55
US-09-918-995-9561
; Sequence 9561, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9561
;   LENGTH: 482
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(482)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-9561

  Query Match             13.7%;  Score 300;  DB 3;  Length 482;
  Best Local Similarity   99.7%;  Pred. No. 8e-151;
  Matches  350;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         106 TCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 165
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         132 TCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 191

Qy         166 AGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATG 225
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         192 AGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATG 251

Qy         226 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 285
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         252 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 311

Qy         286 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 345
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         312 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 371

Qy         346 TGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTC 405
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         372 TCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTC 431

Qy         406 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 456
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         432 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGC 482


RESULT 56
US-10-242-535A-34175
; Sequence 34175, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34175
;   LENGTH: 312
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-34175

  Query Match             13.6%;  Score 299;  DB 7;  Length 312;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-150;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1890 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 1949
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          14 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 73

Qy        1950 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAT 2009
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          74 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAT 133

Qy        2010 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 2069
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         134 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 193

Qy        2070 AACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 2129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         194 AACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 253

Qy        2130 TCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCAT 2188
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         254 TCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCAT 312
```

```
RESULT 57
US-10-085-783A-34175
; Sequence 34175, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34175
;   LENGTH: 312
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-34175

  Query Match           13.6%;  Score 299;  DB 7;  Length 312;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-150;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1890 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 1949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            14 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 73

Qy       1950 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAT 2009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            74 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAT 133

Qy       2010 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 2069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           134 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 193

Qy       2070 AACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 2129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           194 AACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTT 253

Qy       2130 TCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCAT 2188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           254 TCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCAT 312


RESULT 58
US-10-242-535A-21865
; Sequence 21865, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21865
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (22)..(22)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-21865

  Query Match           13.6%;  Score 299;  DB 7;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-150;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1784 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            47 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 106

Qy       1844 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           107 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 166

Qy       1904 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 1963
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           167 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 226

Qy       1964 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           227 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 286

Qy       2024 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 2082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           287 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 345


RESULT 59
US-10-085-783A-21865
; Sequence 21865, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21865
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (22)..(22)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-21865

  Query Match             13.6%;  Score 299;  DB 7;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-150;
  Matches  299;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1784 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 106

Qy       1844 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 166

Qy       1904 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 1963
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGA 226

Qy       1964 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 2023
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 286

Qy       2024 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 2082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 345


RESULT 60
US-09-918-995-3620
; Sequence 3620, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3620
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-3620

  Query Match             13.4%;  Score 294;  DB 3;  Length 393;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-147;
  Matches  294;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          45 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 104

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         105 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 164

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         165 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 224

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         225 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 284

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         285 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 338


RESULT 61
US-10-027-632-3839/c
; Sequence 3839, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3839
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-3839

  Query Match             13.1%;  Score 288;  DB 5;  Length 465;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-144;
  Matches  458;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 406

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 346

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC-AGA 255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        345 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAAGA 286

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 226

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        225 CACTCACTGGCAAGACCATCACCCTTGAGGTSGAGCCCAGTGACACCATCGAGAACGTCA 166

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 106

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 46

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 535
```

```
                 ||||||||||||||||||||||||||||||||||||||||
Db          45 TGCACCTGGTGCTCCGTCTCAGAGGTGGATGCAGATCTT 6

RESULT 62
US-10-027-632-3839/c
; Sequence 3839, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3839
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-3839

  Query Match             13.1%;  Score 288;  DB 6;  Length 465;
  Best Local Similarity   99.6%;  Pred. No. 2.4e-144;
  Matches  458;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         465 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 406

Qy         137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         405 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 346

Qy         197 GATCTTTGCTGGAAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC-AGA 255
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db         345 GATCTTTGCTGGAAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAAGA 286

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         285 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 226

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         225 CACTCACTGGCAAGACCATCACCCTTGAGGTSGAGCCCAGTGACACCATCGAGAACGTCA 166

Qy         376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         165 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 106

Qy         436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         105 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 46

Qy         496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 535
               ||||||||||||||||||||||||||||||||||||||||
Db          45 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 6

RESULT 63
US-10-242-535A-49271
; Sequence 49271, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49271
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-49271

  Query Match             12.8%;  Score 281;  DB 7;  Length 332;
  Best Local Similarity   99.7%;  Pred. No. 1.4e-140;
  Matches  331;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1790 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1849
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 60

Qy        1850 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 1909
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 120

Qy        1910 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 1969
               |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db         121 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCGGTGACACCATCGAGAATGT 180

Qy        1970 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 2029
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 240

Qy        2030 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 2089
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 300

Qy        2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 2121
               ||||||||||||||||||||||||||||||||
Db         301 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 332

RESULT 64
US-10-085-783A-49271
; Sequence 49271, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
```

```
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49271
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-49271

  Query Match             12.8%;  Score 281;  DB 7;  Length 332;
  Best Local Similarity   99.7%;  Pred. No. 1.4e-140;
  Matches  331;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1790 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 1849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 60

Qy       1850 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 1909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 120

Qy       1910 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 1969
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||.
Db        121 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCGGTGACACCATCGAGAATGT 180

Qy       1970 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 2029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 240

Qy       2030 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 2089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 300

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 2121
             ||||||||||||||||||||||||||||||||
Db        301 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTG 332


RESULT 65
US-10-242-535A-31879
; Sequence 31879, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31879
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(32)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-31879

  Query Match             12.6%;  Score 276;  DB 7;  Length 361;
  Best Local Similarity   99.7%;  Pred. No. 7e-138;
  Matches  326;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          33 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 92

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          93 ACCATCACCCTCGAGGTGGAACCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 152

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 2045
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         153 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 212

Qy       2046 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 2105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         213 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 272

Qy       2106 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 2165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         273 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 332

Qy       2166 TCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||
Db         333 TCCTTTCAATAAAGTTGTTGCATTCCC 359


RESULT 66
US-10-085-783A-31879
; Sequence 31879, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31879
;   LENGTH: 361
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(32)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-31879

  Query Match             12.6%;  Score 276;  DB 7;  Length 361;
  Best Local Similarity   99.7%;  Pred. No. 7e-138;
  Matches  326;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          33 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 92

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          93 ACCATCACCCTCGAGGTGGAACCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 152

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 2045
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         153 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 212

Qy       2046 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 2105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         213 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 272
```

```
Qy       2106 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 2165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 332

Qy       2166 TCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||
Db        333 TCCTTTCAATAAAGTTGTTGCATTCCC 359


RESULT 67
US-09-918-995-7727
; Sequence 7727, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7727
;   LENGTH: 399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-7727

  Query Match             12.5%;  Score 275;  DB 3;  Length 399;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-137;
  Matches  275;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 106

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 166

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 226

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 286

Qy        456 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
             |||||||||||||||||||||||||||||||||||
Db        287 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 321


RESULT 68
US-10-242-535A-33670
; Sequence 33670, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33670
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107)..(107)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (390)..(390)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (393)..(393)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-33670

  Query Match             12.4%;  Score 271;  DB 7;  Length 396;
  Best Local Similarity   99.7%;  Pred. No. 3.5e-135;
  Matches  321;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 60

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGNACCCTGTCTGACT 120

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 180

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 240

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 300

Qy       1336 TGATCTTTGCTGGGAAACAGCT 1357
             |||||||||||||||||||||||
Db        301 TGATCTTTGCTGGGAAACAGCT 322


RESULT 69
US-10-085-783A-33670
; Sequence 33670, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33670
;   LENGTH: 396
;   TYPE: DNA
```

```
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (107)..(107)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (390)..(390)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (393)..(393)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-33670

  Query Match            12.4%;  Score 271;  DB 7;  Length 396;
  Best Local Similarity  99.7%;  Pred. No. 3.5e-135;
  Matches  321;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 60

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGNACCCTGTCTGACT 120

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 180

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 240

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 300

Qy       1336 TGATCTTTGCTGGGAAACAGCT 1357
             |||||||||||||||||||||||
Db        301 TGATCTTTGCTGGGAAACAGCT 322


RESULT 70
US-10-004-427-52
; Sequence 52, Application US/10004427
; Publication No. US20030082512A1
; GENERAL INFORMATION:
;  APPLICANT: Yerramilli, Subrahmanyam V.
;  APPLICANT:  Prashar, Yatindra
;  APPLICANT:  Newberger, Peter
;  APPLICANT:  Goguen, Jon
;  APPLICANT:  Weissman, Sherman M.
;  TITLE OF INVENTION: A PROCESS TO STUDY CHANGES IN GENE EXPRESSION IN
;  TITLE OF INVENTION:  GRANULOCYTIC CELLS
;  FILE REFERENCE: 44921-5016-US
;  CURRENT APPLICATION NUMBER: US/10/004,427
;  CURRENT FILING DATE:  2001-12-06
;  PRIOR APPLICATION NUMBER: US/09/506,729
;  PRIOR FILING DATE: 2000-02-18
;  PRIOR APPLICATION NUMBER: PCT/US98/17284
;  PRIOR FILING DATE: 1998-08-21
;  PRIOR APPLICATION NUMBER: 60/056,844
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-004-427-52

  Query Match            12.3%;  Score 270;  DB 5;  Length 284;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-134;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 74

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 134

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 194

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 254

Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTG 2182
             |||||||||||||||||||||||||||||
Db        255 TTTCATTGCACTTTCCTTTCAATAAAGTTG 284


RESULT 71
US-10-984-752-52
; Sequence 52, Application US/10984752
; Publication No. US20050084896A1
; GENERAL INFORMATION:
;  APPLICANT: Yerramilli, Subrahmanyam V.
;  APPLICANT:  Prashar, Yatindra
;  APPLICANT:  Newberger, Peter
;  APPLICANT:  Goguen, Jon
;  APPLICANT:  Weissman, Sherman M.
;  TITLE OF INVENTION: A PROCESS TO STUDY CHANGES IN GENE EXPRESSION IN
;  TITLE OF INVENTION:  GRANULOCYTIC CELLS
;  FILE REFERENCE: 44921-5016-US
;  CURRENT APPLICATION NUMBER: US/10/984,752
;  CURRENT FILING DATE:  2004-11-10
;  PRIOR APPLICATION NUMBER: US/10/984,752
;  PRIOR FILING DATE: 2004-11-10
;  PRIOR APPLICATION NUMBER: PCT/US98/17284
;  PRIOR FILING DATE: 1998-08-21
;  PRIOR APPLICATION NUMBER: 60/056,844
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-984-752-52

  Query Match            12.3%;  Score 270;  DB 9;  Length 284;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-134;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 74

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 134

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 194

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 254
```

```
Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTG 2182
              ||||||||||||||||||||||||||||||
Db        255 TTTCATTGCACTTTCCTTTCAATAAAGTTG 284


RESULT 72
US-10-242-535A-12261
; Sequence 12261, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12261
;   LENGTH: 288
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-12261

  Query Match             12.3%;  Score 270;  DB 7;  Length 288;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-134;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1898 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACAC 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACAC 78

Qy       1958 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAG 2017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAG 138

Qy       2018 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 2077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 198

Qy       2078 GAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTT 2137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTT 258

Qy       2138 TAAGGTTTCAACAAATTTCATTGCACTTTC 2167
              ||||||||||||||||||||||||||||||
Db        259 TAAGGTTTCAACAAATTTCATTGCACTTTC 288


RESULT 73
US-10-085-783A-12261
; Sequence 12261, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12261
;   LENGTH: 288
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-12261

  Query Match             12.3%;  Score 270;  DB 7;  Length 288;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-134;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1898 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACAC 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACAC 78

Qy       1958 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAG 2017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAG 138

Qy       2018 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 2077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 198

Qy       2078 GAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTT 2137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 GAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTT 258

Qy       2138 TAAGGTTTCAACAAATTTCATTGCACTTTC 2167
              ||||||||||||||||||||||||||||||
Db        259 TAAGGTTTCAACAAATTTCATTGCACTTTC 288


RESULT 74
US-09-918-995-8065
; Sequence 8065, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8065
;   LENGTH: 411
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-8065

  Query Match             11.9%;  Score 260;  DB 3;  Length 411;
  Best Local Similarity   100.0%;  Pred. No. 3e-129;
  Matches  260;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 60

Qy         87 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 120

Qy        147 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 180

Qy        207 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 266
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 240

Qy        267 CTGCACCTGGTGCTCCGTCT 286
                ||||||||||||||||||||
Db        241 CTGCACCTGGTGCTCCGTCT 260


RESULT 75
US-10-242-535A-43141
; Sequence 43141, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43141
;   LENGTH: 345
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (265)..(265)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (280)..(280)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (296)..(296)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (328)..(328)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (331)..(331)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-43141

  Query Match           11.8%;  Score 259;  DB 7;  Length 345;
  Best Local Similarity  100.0%;  Pred. No. 1e-128;
  Matches  259;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1854 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1913
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 65

Qy       1914 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 1973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 125

Qy       1974 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 2033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 185

Qy       2034 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 2093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 245

Qy       2094 CACTTGGTCCTGCGCTTGA 2112
                ||||||||||||||||||
Db        246 CACTTGGTCCTGCGCTTGA 264


RESULT 76
US-10-085-783A-43141
; Sequence 43141, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43141
;   LENGTH: 345
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (265)..(265)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (280)..(280)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (296)..(296)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (328)..(328)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (331)..(331)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-43141

  Query Match           11.8%;  Score 259;  DB 7;  Length 345;
  Best Local Similarity  100.0%;  Pred. No. 1e-128;
  Matches  259;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1854 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1913
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 65

Qy       1914 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 1973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 125

Qy       1974 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 2033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 185

Qy       2034 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 2093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 245

Qy       2094 CACTTGGTCCTGCGCTTGA 2112
```

```
                  ||||||||||||||||||
Db        246 CACTTGGTCCTGCGCTTGA 264


RESULT 77
US-10-242-535A-31872
; Sequence 31872, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31872
;   LENGTH: 339
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (25)..(25)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (283)..(283)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-31872

  Query Match             11.8%;  Score 258;  DB 7;  Length 339;
  Best Local Similarity   99.7%;  Pred. No. 3.6e-128;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1400 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 90

Qy       1460 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 150

Qy       1520 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 210

Qy       1580 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 270

Qy       1640 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1699
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CCTGGTGCTCCGNCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 330

Qy       1700 CATCACTCT 1708
             |||||||||
Db        331 CATCACTCT 339


RESULT 78
US-10-085-783A-31872
; Sequence 31872, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31872
;   LENGTH: 339
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (25)..(25)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (283)..(283)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-31872

  Query Match             11.8%;  Score 258;  DB 7;  Length 339;
  Best Local Similarity   99.7%;  Pred. No. 3.6e-128;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1400 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1459
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 90

Qy       1460 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 150

Qy       1520 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1579
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 210

Qy       1580 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 270

Qy       1640 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1699
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CCTGGTGCTCCGNCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 330

Qy       1700 CATCACTCT 1708
             |||||||||
Db        331 CATCACTCT 339


RESULT 79
US-10-242-535A-50008
; Sequence 50008, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
```

```
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50008
;   LENGTH: 345
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-50008

  Query Match            11.8%;  Score 258;  DB 7;  Length 345;
  Best Local Similarity   99.7%;  Pred. No. 3.6e-128;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1834 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 1893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 96

Qy       1894 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTG 1953
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 TGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTG 156

Qy       1954 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGC 2013
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGC 216

Qy       2014 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 2073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 276

Qy       2074 TCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCC 2133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCC 336

Qy       2134 CTTTTAAGG 2142
             |||||||||
Db        337 CTTTTAAGG 345


RESULT 80
US-10-085-783A-50008
; Sequence 50008, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50008
;   LENGTH: 345
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-50008

  Query Match            11.8%;  Score 258;  DB 7;  Length 345;
  Best Local Similarity   99.7%;  Pred. No. 3.6e-128;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1834 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 1893
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 96

Qy       1894 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTG 1953
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 TGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTG 156

Qy       1954 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGC 2013
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGC 216

Qy       2014 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 2073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 276

Qy       2074 TCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCC 2133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCC 336

Qy       2134 CTTTTAAGG 2142
             |||||||||
Db        337 CTTTTAAGG 345


RESULT 81
US-10-242-535A-30623
; Sequence 30623, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30623
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28)..(28)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (288)..(288)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-30623

  Query Match            11.6%;  Score 255;  DB 7;  Length 337;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-126;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1332 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1391
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 92

Qy       1392 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 1451
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 152

Qy       1452 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1511
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 212

Qy       1512 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1571
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 272

Qy       1572 GCTGGGAAACAGCTG 1586
                 |||||||||||||||
Db        273 GCTGGGAAACAGCTG 287


RESULT 82
US-10-085-783A-30623
; Sequence 30623, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30623
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28)..(28)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (288)..(288)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-30623

  Query Match           11.6%;  Score 255;  DB 7;  Length 337;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-126;
  Matches  255;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1332 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1391
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 92

Qy       1392 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 1451
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 152

Qy       1452 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1511
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 212

Qy       1512 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1571
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 272

Qy       1572 GCTGGGAAACAGCTG 1586
                 |||||||||||||||
Db        273 GCTGGGAAACAGCTG 287


RESULT 83
US-09-884-441-45
; Sequence 45, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C7
;  CURRENT APPLICATION NUMBER: US/09/884,441
;  CURRENT FILING DATE:  2001-06-18
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-884-441-45

  Query Match           11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity  100.0%;  Pred. No. 8.9e-123;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
                 ||||||||
Db        532 AAGGTTTC 539


RESULT 84
US-09-907-969-45
; Sequence 45, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Hill, Paul
;  APPLICANT:  Albone, Earl
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C8
```

```
;  CURRENT APPLICATION NUMBER: US/09/907,969
;  CURRENT FILING DATE:  2001-07-17
;  NUMBER OF SEQ ID NOS: 596
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-907-969-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-123;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 85
US-09-827-271-45
; Sequence 45, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-123;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 86
US-10-198-053-45
; Sequence 45, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-45

  Query Match             11.3%;  Score 248;  DB 6;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-123;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 87
US-10-860-790-45
; Sequence 45, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C11
;  CURRENT APPLICATION NUMBER: US/10/860,790
;  CURRENT FILING DATE:  2004-06-02
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-10-860-790-45

  Query Match            11.3%;  Score 248;  DB 8;  Length 585;
  Best Local Similarity  100.0%;  Pred. No. 8.9e-123;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 88
US-10-357-930-45895/c
; Sequence 45895, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45895
;   LENGTH: 467
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-357-930-45895

  Query Match            11.1%;  Score 244;  DB 8;  Length 467;
  Best Local Similarity  99.4%;  Pred. No. 1.3e-120;
  Matches  344;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        343 AGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCA 402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCA 408

Qy        403 TTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TTCCTCTTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 348

Qy        463 TGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TGTCTGACTACAACATCCAGAAAGAGTTTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 288

Qy        523 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCA 228

Qy        583 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATC 642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATC 168

Qy        643 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        167 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 122


RESULT 89
US-10-242-535A-10811
; Sequence 10811, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10811
;   LENGTH: 329
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (243)..(243)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (251)..(251)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-10811

  Query Match            11.0%;  Score 242;  DB 7;  Length 329;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-119;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1411 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 60

Qy       1471 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1530
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 120

Qy       1531 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 1590
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 180

Qy       1591 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 1650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 240

Qy       1651 GT 1652
              ||
Db        241 GT 242


RESULT 90
US-10-085-783A-10811
; Sequence 10811, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10811
;   LENGTH: 329
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (243)..(243)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (251)..(251)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-10811

  Query Match            11.0%;  Score 242;  DB 7;  Length 329;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-119;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1411 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 60

Qy       1471 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 120

Qy       1531 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 1590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 180

Qy       1591 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 1650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 240

Qy       1651 GT 1652
              ||
Db        241 GT 242


RESULT 91
US-10-242-535A-51059
; Sequence 51059, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51059
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(33)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-51059

  Query Match            10.9%;  Score 238;  DB 7;  Length 337;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-117;
  Matches  238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 136

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 196

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 2134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 256

Qy       2135 TTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 314


RESULT 92
US-10-085-783A-51059
; Sequence 51059, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51059
;   LENGTH: 337
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(33)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-51059
```

```
  Query Match            10.9%;  Score 238;  DB 7;  Length 337;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-117;
  Matches  238;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 136

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 196

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 2134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 256

Qy       2135 TTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 TTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 314


RESULT 93
US-10-242-535A-28641
; Sequence 28641, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28641
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (413)..(413)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (417)..(417)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (457)..(457)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (473)..(473)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-28641

  Query Match            10.8%;  Score 236;  DB 7;  Length 474;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-116;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 95

Qy       1575 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 155

Qy       1635 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 215

Qy       1695 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 271


RESULT 94
US-10-085-783A-28641
; Sequence 28641, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28641
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (413)..(413)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (417)..(417)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (457)..(457)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (473)..(473)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-28641

  Query Match            10.8%;  Score 236;  DB 7;  Length 474;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-116;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 95

Qy       1575 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 155

Qy       1635 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 215

Qy       1695 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 271
```

```
RESULT 95
US-10-242-535A-33237
; Sequence 33237, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33237
;   LENGTH: 286
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-33237

  Query Match            10.7%;  Score 235;  DB 7;  Length 286;
  Best Local Similarity   99.7%;  Pred. No. 9.2e-116;
  Matches  285;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1337 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 60

Qy       1397 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 120

Qy       1457 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 180

Qy       1517 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 240

Qy       1577 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             |||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 286


RESULT 96
US-10-085-783A-33237
; Sequence 33237, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33237
;   LENGTH: 286
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-33237

  Query Match            10.7%;  Score 235;  DB 7;  Length 286;
  Best Local Similarity   99.7%;  Pred. No. 9.2e-116;
  Matches  285;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1337 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 60

Qy       1397 AGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 1456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGAC 120

Qy       1457 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 1516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAA 180

Qy       1517 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 1576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGG 240

Qy       1577 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             |||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 286


RESULT 97
US-10-242-535A-27365
; Sequence 27365, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27365
;   LENGTH: 237
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-27365

  Query Match            10.6%;  Score 233;  DB 7;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-114;
  Matches  233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1960 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 2019
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 64

Qy       2020 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 2079
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 124
```

```
Qy       2080 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 2139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 184

Qy       2140 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 237


RESULT 98
US-10-085-783A-27365
; Sequence 27365, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27365
;   LENGTH: 237
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-27365

  Query Match           10.6%;  Score 233;  DB 7;  Length 237;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-114;
  Matches  233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1960 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 2019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 64

Qy       2020 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 2079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 124

Qy       2080 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 2139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 184

Qy       2140 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 237


RESULT 99
US-10-242-535A-34542
; Sequence 34542, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34542
;   LENGTH: 232
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-34542

  Query Match           10.6%;  Score 232;  DB 7;  Length 232;
  Best Local Similarity  100.0%;  Pred. No. 3.8e-114;
  Matches  232;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 232


RESULT 100
US-10-085-783A-34542
; Sequence 34542, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34542
;   LENGTH: 232
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-34542

  Query Match           10.6%;  Score 232;  DB 7;  Length 232;
  Best Local Similarity  100.0%;  Pred. No. 3.8e-114;
  Matches  232;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 60

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 120

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 180

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 232


RESULT 101
US-09-796-692-3431/c
; Sequence 3431, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3431
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-3431

  Query Match             10.4%;  Score 228;  DB 3;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-112;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 169

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 109

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 49

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 1


RESULT 102
US-10-040-862-3431/c
; Sequence 3431, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/040,862
;  CURRENT FILING DATE:  2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: US 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: US 09/796,692
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3431
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-040-862-3431

  Query Match             10.4%;  Score 228;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-112;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 169

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 109

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 49

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 1


RESULT 103
US-10-057-475B-3431/c
; Sequence 3431, Application US/10057475B
```

```
; Publication No. US20040002068A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Clapper, Jonathan David
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Ordonez, Nadia
;  APPLICANT:  Carter, Lauren
;  APPLICANT:  McNeill, Patricia Dianne
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-014402US
;  CURRENT APPLICATION NUMBER: US/10/057,475B
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10979
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3431
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-057-475B-3431

  Query Match             10.4%;  Score 228;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-112;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 169

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 109

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 49

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 1


RESULT 104
US-10-154-884B-3431/c
; Sequence 3431, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Thera
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013521US
;  CURRENT APPLICATION NUMBER: US/10/154,884B
;  CURRENT FILING DATE:  2002-05-23
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 11290
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3431
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-154-884B-3431

  Query Match             10.4%;  Score 228;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-112;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 169

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 109

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 49

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 1


RESULT 105
US-10-764-324-3431/c
; Sequence 3431, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/764,324
;  CURRENT FILING DATE:  2004-01-23
```

```
;  PRIOR APPLICATION NUMBER: US/10/040,862
;  PRIOR FILING DATE: 2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3431
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-764-324-3431

  Query Match             10.4%;  Score 228;  DB 8;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-112;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        216 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 169

Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 109

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 49

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              |||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 1


RESULT 106
US-10-242-535A-49909
; Sequence 49909, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49909
;   LENGTH: 232
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-49909

  Query Match             10.2%;  Score 223;  DB 7;  Length 232;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-109;
  Matches  223;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1970 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 2029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 60

Qy       2030 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 2089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 120

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 2149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 180

Qy       2150 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 223


RESULT 107
US-10-085-783A-49909
; Sequence 49909, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49909
;   LENGTH: 232
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-49909

  Query Match             10.2%;  Score 223;  DB 7;  Length 232;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-109;
  Matches  223;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1970 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 2029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 60

Qy       2030 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 2089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 120

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 2149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 180

Qy       2150 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 223


RESULT 108
US-10-242-535A-42741
```

; Sequence 42741, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42741
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-42741

Query Match 9.9%; Score 218; DB 7; Length 403;
Best Local Similarity 99.2%; Pred. No. 1.3e-106;
Matches 368; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1560 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 92

Qy       1620 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 1679
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         93 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 152

Qy       1680 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1739
             |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        153 AAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATCGAGAAT 212

Qy       1740 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 272

Qy       1800 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 332

Qy       1860 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 392

Qy       1920 GGTAAGACCAT 1930
             ||||||||||||
Db        393 GGTAAGACCAT 403
```

RESULT 109
US-10-085-783A-42741
; Sequence 42741, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42741
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-42741

Query Match 9.9%; Score 218; DB 7; Length 403;
Best Local Similarity 99.2%; Pred. No. 1.3e-106;
Matches 368; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1560 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 92

Qy       1620 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTG 1679
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         93 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 152

Qy       1680 AAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1739
             |||||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        153 AAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATCGAGAAT 212

Qy       1740 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 272

Qy       1800 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1859
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 332

Qy       1860 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1919
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 392

Qy       1920 GGTAAGACCAT 1930
             ||||||||||||
Db        393 GGTAAGACCAT 403
```

RESULT 110
US-10-425-115-23758
; Sequence 23758, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 23758
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_121671C.1
US-10-425-115-23758

```
  Query Match             9.9%;  Score 217;  DB 8;  Length 299;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-106;
  Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1469 GACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA 1528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA 127

Qy       1529 AGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 1588
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 AGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 187

Qy       1589 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCT 1648
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCT 247

Qy       1649 CCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
             |||||||||||||||||||||||||||||||||||||
Db        248 CCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 284


RESULT 111
US-10-242-535A-25883
; Sequence 25883, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25883
;   LENGTH: 372
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (38)..(38)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (108)..(108)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (110)..(110)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (118)..(118)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-25883

  Query Match             9.9%;  Score 217;  DB 7;  Length 372;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-106;
  Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 215

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 275

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 335

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             |||||||||||||||||||||||||||||||||||||
Db        336 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 372


RESULT 112
US-10-085-783A-25883
; Sequence 25883, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25883
;   LENGTH: 372
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (38)..(38)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (108)..(108)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (110)..(110)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (118)..(118)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-25883

  Query Match             9.9%;  Score 217;  DB 7;  Length 372;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-106;
  Matches  217;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 215

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 275
```

```
Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 335

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
             |||||||||||||||||||||||||||||||||||||
Db        336 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 372


RESULT 113
US-10-242-535A-27257
; Sequence 27257, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27257
;   LENGTH: 249
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(8)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (33)..(34)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-27257

  Query Match             9.7%;  Score 213;  DB 7;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-104;
  Matches  213;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1425 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 1484
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 96

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 156

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 216

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             |||||||||||||||||||||||||||||||||
Db        217 TCTGACTACAACATCCAGAAAGAGTCCACCCTG 249


RESULT 114
US-10-085-783A-27257
; Sequence 27257, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27257
;   LENGTH: 249
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(8)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (33)..(34)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-27257

  Query Match             9.7%;  Score 213;  DB 7;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-104;
  Matches  213;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1425 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 1484
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 CTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAA 96

Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 156

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 216

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
             |||||||||||||||||||||||||||||||||
Db        217 TCTGACTACAACATCCAGAAAGAGTCCACCCTG 249


RESULT 115
US-09-876-889-277
; Sequence 277, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
;  TITLE OF INVENTION:  CANCER THERAPY AND DIAGNOSIS
;  FILE REFERENCE: 210121.466C3
;  CURRENT APPLICATION NUMBER: US/09/876,889
;  CURRENT FILING DATE:  2001-06-06
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 277
;   LENGTH: 440
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(440)
;   OTHER INFORMATION: n = A,T,C or G
US-09-876-889-277

  Query Match             9.6%;  Score 211;  DB 3;  Length 440;
  Best Local Similarity   100.0%;  Pred. No. 8e-103;
  Matches  211;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 60

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 966
              ||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 211


RESULT 116
US-10-242-535A-30987
; Sequence 30987, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30987
;   LENGTH: 261
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-30987

  Query Match             9.6%;  Score 210;  DB 7;  Length 261;
  Best Local Similarity   99.6%;  Pred. No. 2.8e-102;
  Matches  260;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1130 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 60

Qy       1190 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 120

Qy       1250 TCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1309
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 TCTCGAGGTGGAGCCGAGTGACCCCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 180

Qy       1310 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 240

Qy       1370 CACCCTGTCTGACTACAACAT 1390
              |||||||||||||||||||||
Db        241 CACCCTGTCTGACTACAACAT 261


RESULT 117
US-10-085-783A-30987
; Sequence 30987, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30987
;   LENGTH: 261
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-30987

  Query Match             9.6%;  Score 210;  DB 7;  Length 261;
  Best Local Similarity   99.6%;  Pred. No. 2.8e-102;
  Matches  260;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1130 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 60

Qy       1190 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 120

Qy       1250 TCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1309
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        121 TCTCGAGGTGGAGCCGAGTGACCCCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 180

Qy       1310 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 240

Qy       1370 CACCCTGTCTGACTACAACAT 1390
              |||||||||||||||||||||
Db        241 CACCCTGTCTGACTACAACAT 261


RESULT 118
US-10-242-535A-8262
; Sequence 8262, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
```

; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8262
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (252)..(252)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (298)..(298)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309)..(309)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (318)..(318)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-8262

Query Match 9.6%; Score 210; DB 7; Length 330;
Best Local Similarity 100.0%; Pred. No. 2.8e-102;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 101

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 161

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 221

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGG 226
              ||||||||||||||||||||||||||||||
Db        222 GATCTTTGCTGGAAAACAGCTGGAAGATGG 251
```

RESULT 119
US-10-085-783A-8262
; Sequence 8262, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8262
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (252)..(252)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (298)..(298)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309)..(309)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (318)..(318)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-8262

Query Match 9.6%; Score 210; DB 7; Length 330;
Best Local Similarity 100.0%; Pred. No. 2.8e-102;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 101

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 161

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 221

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGG 226
              ||||||||||||||||||||||||||||||
Db        222 GATCTTTGCTGGAAAACAGCTGGAAGATGG 251
```

RESULT 120
US-09-938-864-384
; Sequence 384, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Smithgall, Molly
; APPLICANT: Moulton, Gus
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally
; APPLICANT: Evans, Lawrence
; APPLICANT: Spies, A. Gregory
; APPLICANT: Boydston, Jeremy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C5
; CURRENT APPLICATION NUMBER: US/09/938,864

```
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-938-864-384

  Query Match             9.4%;  Score 207;  DB 3;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 121
US-10-125-635A-384
; Sequence 384, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-125-635A-384

  Query Match             9.4%;  Score 207;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 122
US-10-121-988-80
; Sequence 80, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
;  APPLICANT: Hosken, Nancy Ann
;  APPLICANT:  McGowan, Patrick
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally P.
;  APPLICANT:  Evans, Lawrence S.
;  APPLICANT:  Swanson, Ryan M.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
;  FILE REFERENCE: 210121.538C1
;  CURRENT APPLICATION NUMBER: US/10/121,988
;  CURRENT FILING DATE:  2002-04-11
;  NUMBER OF SEQ ID NOS: 183
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens-ubiquitin UL49-HSV-2
US-10-121-988-80

  Query Match             9.4%;  Score 207;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 123
US-10-002-603-384
; Sequence 384, Application US/10002603
; Publication No. US20030095971A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C6
;  CURRENT APPLICATION NUMBER: US/10/002,603
;  CURRENT FILING DATE:  2001-10-30
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
;   LENGTH: 228
;   TYPE: DNA
```

```
;   ORGANISM: Homo sapiens
US-10-002-603-384

  Query Match             9.4%;  Score 207;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 124
US-10-200-562-80
; Sequence 80, Application US/10200562
; Publication No. US20030165819A1
; GENERAL INFORMATION:
;  APPLICANT: McGowen, Patrick
;  APPLICANT:  Hosken, Nancy A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
;  FILE REFERENCE: 210121.538C2
;  CURRENT APPLICATION NUMBER: US/10/200,562
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens-ubiquitin UL49-HSV-2
US-10-200-562-80

  Query Match             9.4%;  Score 207;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 125
US-10-237-551-80
; Sequence 80, Application US/10237551
; Publication No. US20030165820A1
; GENERAL INFORMATION:
;  APPLICANT: Day, Craig H.
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Parsons, Joseph M.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
;  FILE REFERENCE: 210121.538C3
;  CURRENT APPLICATION NUMBER: US/10/237,551
;  CURRENT FILING DATE:  2002-09-06
;  NUMBER OF SEQ ID NOS: 254
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens-ubiquitin UL49-HSV-2
US-10-237-551-80

  Query Match             9.4%;  Score 207;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 126
US-10-195-835-384
; Sequence 384, Application US/10195835
; Publication No. US20030198622A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  APPLICANT:  Mossman, Sally P.
;  APPLICANT:  Evans, Lawrence S.
;  APPLICANT:  Swanson, Ryan M.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C8
;  CURRENT APPLICATION NUMBER: US/10/195,835
;  CURRENT FILING DATE:  2002-07-12
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-195-835-384

  Query Match             9.4%;  Score 207;  DB 6;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141
```

```
Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228
```

RESULT 127
US-10-286-333-384
; Sequence 384, Application US/10286333
; Publication No. US20030215458A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Jaya, No. US20030215458A1alie
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C10
; CURRENT APPLICATION NUMBER: US/10/286,333
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-286-333-384

Query Match 9.4%; Score 207; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-100;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228
```

RESULT 128
US-10-244-830-384
; Sequence 384, Application US/10244830
; Publication No. US20030235557A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C9
; CURRENT APPLICATION NUMBER: US/10/244,830
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-244-830-384

Query Match 9.4%; Score 207; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-100;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228
```

RESULT 129
US-10-427-717-384
; Sequence 384, Application US/10427717
; Publication No. US20040018204A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Jaya, No. US20040018204A1alie
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
; TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
; FILE REFERENCE: 210121.465C11
; CURRENT APPLICATION NUMBER: US/10/427,717
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 508
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-427-717-384

Query Match 9.4%; Score 207; DB 7; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.2e-100;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228
```

RESULT 130
US-10-648-780-384
; Sequence 384, Application US/10648780
; Publication No. US20040126362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Smithgall, Molly D.
; APPLICANT: Carter, Darrick
; APPLICANT: Cheever, Martin A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Sutherland, R. Alec
; APPLICANT: Mossman, Sally P.

```
;  APPLICANT:  Evans, Lawrence S.
;  APPLICANT:  Swanson, Ryan M.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C8
;  CURRENT APPLICATION NUMBER: US/10/648,780
;  CURRENT FILING DATE:  2003-08-26
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 384
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-648-780-384

  Query Match             9.4%;  Score 207;  DB 7;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 131
US-10-945-050-80
; Sequence 80, Application US/10945050
; Publication No. US20050130132A1
; GENERAL INFORMATION:
;  APPLICANT: Day, Craig H.
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Parsons, Joseph M.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
;  FILE REFERENCE: 538C4
;  CURRENT APPLICATION NUMBER: US/10/945,050
;  CURRENT FILING DATE:  2004-09-20
;  NUMBER OF SEQ ID NOS: 254
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens-ubiquitin UL49-HSV-2
US-10-945-050-80

  Query Match             9.4%;  Score 207;  DB 9;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 132
US-10-475-203A-1
; Sequence 1, Application US/10475203A
; Publication No. US20040241177A1
; GENERAL INFORMATION:
;  APPLICANT: FRAZER, Ian
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND USES THEREFOR
;  FILE REFERENCE: E1679-00002
;  CURRENT APPLICATION NUMBER: US/10/475,203A
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: PCT/AU02/00486
;  PRIOR FILING DATE: 2002-04-18
;  PRIOR APPLICATION NUMBER: AU PR4468/01
;  PRIOR FILING DATE: 2001-04-18
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 231
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(231)
;   OTHER INFORMATION:
US-10-475-203A-1

  Query Match             9.4%;  Score 207;  DB 8;  Length 231;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              |||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 133
US-09-871-161-169/c
; Sequence 169, Application US/09871161
; Publication No. US20030097666A1
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION:  PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/871,161
;  CURRENT FILING DATE:  2001-05-31
;  PRIOR APPLICATION NUMBER: 09/328,111
;  PRIOR FILING DATE: 1999-06-08
;  PRIOR APPLICATION NUMBER: 60/117,393
;  PRIOR FILING DATE: 1999-01-27
;  PRIOR APPLICATION NUMBER: 60/098,639
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 169
;   LENGTH: 609
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-871-161-169

  Query Match             9.4%;  Score 207;  DB 3;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-100;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 305

Qy       1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 245

Qy       1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 185

Qy       1871 GGTGCTCCGTCTCAGAGGTGGGATGCA 1897
             ||||||||||||||||||||||||||||
Db        184 GGTGCTCCGTCTCAGAGGTGGGATGCA 158


RESULT 134
US-09-925-299-269
; Sequence 269, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 269
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (19)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (240)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (275)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (296)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (321)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (332)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-269

  Query Match             9.4%;  Score 205;  DB 3;  Length 332;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-99;
  Matches  205;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 94

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 154

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 214

Qy        229 GCACCCTGTCTGACTACAACATCCA 253
              |||||||||||||||||||||||||
Db        215 GCACCCTGTCTGACTACAACATCCA 239


RESULT 135
US-09-925-299-269
; Sequence 269, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA102
;  CURRENT APPLICATION NUMBER: US/09/925,299
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05883
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 269
;   LENGTH: 332
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (19)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (240)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (275)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (296)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (321)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (332)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-269

  Query Match             9.4%;  Score 205;  DB 3;  Length 332;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-99;
  Matches  205;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 94

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 154

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        155 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 214

Qy        229 GCACCCTGTCTGACTACAACATCCA 253
              |||||||||||||||||||||||||
Db        215 GCACCCTGTCTGACTACAACATCCA 239


RESULT 136
US-10-242-535A-3777
; Sequence 3777, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3777
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-3777

  Query Match             9.3%;  Score 204;  DB 7;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-99;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1700 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 60

Qy       1760 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 120

Qy       1820 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 180

Qy       1880 TCTCAGAGGTGGGATGCAGATCTT 1903
              |||||||||||||||||||||||||
Db        181 TCTCAGAGGTGGGATGCAGATCTT 204


RESULT 137
US-10-085-783A-3777
; Sequence 3777, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3777
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-3777

  Query Match             9.3%;  Score 204;  DB 7;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-99;
  Matches  204;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1700 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 60

Qy       1760 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 120

Qy       1820 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 180

Qy       1880 TCTCAGAGGTGGGATGCAGATCTT 1903
              |||||||||||||||||||||||||
Db        181 TCTCAGAGGTGGGATGCAGATCTT 204


RESULT 138
US-10-360-053-9
; Sequence 9, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
;  APPLICANT: Caterer, Nigel
;  APPLICANT:  Uttenthal, Lars O
;  APPLICANT:  Nielsen, Rasmus W
;  TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Anti
;  TITLE OF INVENTION:  Fragments via Antiparallel Hetergeneous Coiled-Coil Peptide Re
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: IMX-0028
;  CURRENT APPLICATION NUMBER: US/10/360,053
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US60/354,376
;  PRIOR FILING DATE: 2002-02-05
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 480
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-10-360-053-9

  Query Match             9.3%;  Score 203;  DB 6;  Length 480;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-98;
  Matches  203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 156

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 216

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 276

Qy        726 CACCTGGTACTCCGTCTCAGAGG 748
              |||||||||||||||||||||||
Db        277 CACCTGGTACTCCGTCTCAGAGG 299
```

```
RESULT 139
US-10-360-053-11
; Sequence 11, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
;  APPLICANT: Caterer, Nigel
;  APPLICANT:  Uttenthal, Lars O
;  APPLICANT:  Nielsen, Rasmus W
;  TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibo
;  TITLE OF INVENTION:  Fragments via Antiparallel Hetergeneous Coiled-Coil Peptide Regi
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE: IMX-0028
;  CURRENT APPLICATION NUMBER: US/10/360,053
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US60/354,376
;  PRIOR FILING DATE: 2002-02-05
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 504
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-10-360-053-11

  Query Match             9.3%;  Score 203;  DB 6;  Length 504;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-98;
  Matches  203;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 180

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 240

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 300

Qy        726 CACCTGGTACTCCGTCTCAGAGG 748
              |||||||||||||||||||||||
Db        301 CACCTGGTACTCCGTCTCAGAGG 323


RESULT 140
US-10-242-535A-36466
; Sequence 36466, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36466
;   LENGTH: 261
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-36466

  Query Match             9.2%;  Score 202;  DB 7;  Length 261;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-98;
  Matches  202;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1989 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 117

Qy       2049 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 177

Qy       2109 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCC 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCC 237

Qy       2169 TTTCAATAAAGTTGTTGCATTC 2190
              |||||||||||||||||||||||
Db        238 TTTCAATAAAGTTGTTGCATTC 259


RESULT 141
US-10-085-783A-36466
; Sequence 36466, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36466
;   LENGTH: 261
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-36466

  Query Match             9.2%;  Score 202;  DB 7;  Length 261;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-98;
  Matches  202;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1989 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 AAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 117

Qy       2049 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGC 177

Qy       2109 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCC 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 TTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCC 237

Qy       2169 TTTCAATAAAGTTGTTGCATTC 2190
              |||||||||||||||||||||||
Db        238 TTTCAATAAAGTTGTTGCATTC 259


RESULT 142
US-10-242-535A-40075
; Sequence 40075, Application US/10242535A
```

```
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40075
;   LENGTH: 490
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(32)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (34)..(34)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (59)..(59)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (212)..(212)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-40075

  Query Match             9.2%;  Score 202;  DB 7;  Length 490;
  Best Local Similarity   99.3%;  Pred. No. 5.7e-98;
  Matches  302;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1787 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 1846
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 169

Qy       1847 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGT 1906
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        170 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGNGGGATGCAGATCTTCGT 229

Qy       1907 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 1966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 289

Qy       1967 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 2026
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        290 TGTCAAGGCAAAGATCCAAGATAAGGAGGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 349

Qy       2027 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 2086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 409

Qy       2087 CACT 2090
             ||||
Db        410 CACT 413


RESULT 143
US-10-085-783A-40075
; Sequence 40075, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40075
;   LENGTH: 490
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(32)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (34)..(34)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (59)..(59)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (212)..(212)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-40075

  Query Match             9.2%;  Score 202;  DB 7;  Length 490;
  Best Local Similarity   99.3%;  Pred. No. 5.7e-98;
  Matches  302;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1787 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 1846
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 169

Qy       1847 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGT 1906
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        170 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGNGGGATGCAGATCTTCGT 229

Qy       1907 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 1966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAA 289

Qy       1967 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 2026
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        290 TGTCAAGGCAAAGATCCAAGATAAGGAGGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 349

Qy       2027 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 2086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 409

Qy       2087 CACT 2090
             ||||
Db        410 CACT 413


RESULT 144
US-09-796-692-3812/c
; Sequence 3812, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
```

```
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3812
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-796-692-3812

  Query Match             9.2%;  Score 201;  DB 3;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 2e-97;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 142

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 82

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 22

Qy        195 CTGATCTTTGCTGGAAAACAG 215
              |||||||||||||||||||||
Db         21 CTGATCTTTGCTGGAAAACAG 1


RESULT 145
US-10-040-862-3812/c
; Sequence 3812, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/040,862
;  CURRENT FILING DATE:  2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: US 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: US 09/796,692
;  PRIOR FILING DATE: 2001-03-01
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3812
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-040-862-3812

  Query Match             9.2%;  Score 201;  DB 5;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 2e-97;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 142

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 82

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 22

Qy        195 CTGATCTTTGCTGGAAAACAG 215
              |||||||||||||||||||||
Db         21 CTGATCTTTGCTGGAAAACAG 1


RESULT 146
US-10-057-475B-3812/c
; Sequence 3812, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Clapper, Jonathan David
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Ordonez, Nadia
;  APPLICANT:  Carter, Lauren
;  APPLICANT:  McNeill, Patricia Dianne
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-014402US
```

```
;  CURRENT APPLICATION NUMBER: US/10/057,475B
;  CURRENT FILING DATE:  2002-01-22
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10979
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3812
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-057-475B-3812

  Query Match             9.2%;  Score 201;  DB 6;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 2e-97;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 142

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 82

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 22

Qy        195 CTGATCTTTGCTGGAAAACAG 215
              |||||||||||||||||||||
Db         21 CTGATCTTTGCTGGAAAACAG 1


RESULT 147
US-10-154-884B-3812/c
; Sequence 3812, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013521US
;  CURRENT APPLICATION NUMBER: US/10/154,884B
;  CURRENT FILING DATE:  2002-05-23
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: US 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 11290
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3812
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-154-884B-3812

  Query Match             9.2%;  Score 201;  DB 6;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 2e-97;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 142

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 82

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 22

Qy        195 CTGATCTTTGCTGGAAAACAG 215
              |||||||||||||||||||||
Db         21 CTGATCTTTGCTGGAAAACAG 1


RESULT 148
US-10-764-324-3812/c
; Sequence 3812, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
;  TITLE OF INVENTION:  Hematological Malignancies
;  FILE REFERENCE: 014058-013520US
;  CURRENT APPLICATION NUMBER: US/10/764,324
;  CURRENT FILING DATE:  2004-01-23
;  PRIOR APPLICATION NUMBER: US/10/040,862
;  PRIOR FILING DATE: 2001-11-06
;  PRIOR APPLICATION NUMBER: US 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: US 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: US 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: US 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: US 60/200,999
;  PRIOR FILING DATE: 2000-05-01
```

```
;  PRIOR APPLICATION NUMBER: US 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: US 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: US 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 10467
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3812
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-764-324-3812

  Query Match             9.2%;  Score 201;  DB 8;  Length 201;
  Best Local Similarity   100.0%;  Pred. No. 2e-97;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 142

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 82

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 22

Qy        195 CTGATCTTTGCTGGAAAACAG 215
              |||||||||||||||||||||
Db         21 CTGATCTTTGCTGGAAAACAG 1


RESULT 149
US-09-918-995-22118
; Sequence 22118, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22118
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(471)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22118

  Query Match             9.2%;  Score 201;  DB 3;  Length 471;
  Best Local Similarity   100.0%;  Pred. No. 2e-97;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 109

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 169

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 229

Qy        229 GCACCCTGTCTGACTACAACA 249
              ||||||||||||||||||||
Db        230 GCACCCTGTCTGACTACAACA 250


RESULT 150
US-10-242-535A-35923
; Sequence 35923, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35923
;   LENGTH: 216
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-35923

  Query Match             9.1%;  Score 200;  DB 7;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-97;
  Matches  200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 76

Qy       2053 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 136

Qy       2113 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 196

Qy       2173 AATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||
Db        197 AATAAAGTTGTTGCATTCCC 216


RESULT 151
US-10-085-783A-35923
; Sequence 35923, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
```

; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35923
; LENGTH: 216
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-35923

Query Match 9.1%; Score 200; DB 7; Length 216;
Best Local Similarity 100.0%; Pred. No. 6.9e-97;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 2052
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 17 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 76

Qy 2053 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 77 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGA 136

Qy 2113 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 2172
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 137 GGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTC 196

Qy 2173 AATAAAGTTGTTGCATTCCC 2192
||||||||||||||||||||
Db 197 AATAAAGTTGTTGCATTCCC 216

RESULT 152
US-10-242-535A-13820
; Sequence 13820, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13820
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (43)..(43)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (67)..(68)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (70)..(71)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (94)..(94)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (295)..(295)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (299)..(299)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (327)..(327)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (356)..(356)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-13820

Query Match 9.1%; Score 200; DB 7; Length 367;
Best Local Similarity 100.0%; Pred. No. 6.9e-97;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 95 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 154

Qy 1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 155 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 214

Qy 1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 274

Qy 1689 ACTGGTAAGACCATCACTCT 1708
||||||||||||||||||||
Db 275 ACTGGTAAGACCATCACTCT 294

RESULT 153
US-10-085-783A-13820
; Sequence 13820, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13820

```
;   LENGTH: 367
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (43)..(43)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (45)..(45)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (67)..(68)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (70)..(71)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (94)..(94)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (295)..(295)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (299)..(299)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (327)..(327)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (356)..(356)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-13820

  Query Match             9.1%;  Score 200;  DB 7;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-97;
  Matches  200;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 154

Qy       1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 214

Qy       1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 274

Qy       1689 ACTGGTAAGACCATCACTCT 1708
             ||||||||||||||||||||
Db        275 ACTGGTAAGACCATCACTCT 294


RESULT 154
US-10-242-535A-50750
; Sequence 50750, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50750
;   LENGTH: 294
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (108)..(108)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (145)..(145)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-50750

  Query Match             8.8%;  Score 192;  DB 7;  Length 294;
  Best Local Similarity   99.3%;  Pred. No. 1.4e-92;
  Matches  292;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 60

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCNGATCAGCAGAGG 120

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 TTGATCTTTGCTGGGAAACAGCTGNAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 240

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 294


RESULT 155
US-10-085-783A-50750
; Sequence 50750, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50750
```

; LENGTH: 294
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (108)..(108)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145)..(145)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-50750

Query Match 8.8%; Score 192; DB 7; Length 294;
Best Local Similarity 99.3%; Pred. No. 1.4e-92;
Matches 292; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 60

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCNGATCAGCAGAGG 120

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 TTGATCTTTGCTGGGAAACAGCTGNAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 240

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 294
```

RESULT 156
US-09-884-441-95/c
; Sequence 95, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; CURRENT FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-884-441-95

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 4.9e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 157
US-09-907-969-95/c
; Sequence 95, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Hill, Paul
; APPLICANT: Albone, Earl
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C8
; CURRENT APPLICATION NUMBER: US/09/907,969
; CURRENT FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 596
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-969-95

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 4.9e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 158
US-09-827-271-95/c
; Sequence 95, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 605
; TYPE: DNA

```
;   ORGANISM: Homo sapien
US-09-827-271-95

  Query Match             8.7%;  Score 191;  DB 3;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 159
US-10-198-053-95/c
; Sequence 95, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-95

  Query Match             8.7%;  Score 191;  DB 6;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 160
US-10-860-790-95/c
; Sequence 95, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C11
;  CURRENT APPLICATION NUMBER: US/10/860,790
;  CURRENT FILING DATE:  2004-06-02
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-860-790-95

  Query Match             8.7%;  Score 191;  DB 8;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 161
US-09-884-441-100
; Sequence 100, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C7
;  CURRENT APPLICATION NUMBER: US/09/884,441
;  CURRENT FILING DATE:  2001-06-18
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-884-441-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521
```

```
Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 162
US-09-907-969-100
; Sequence 100, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Hill, Paul
;  APPLICANT:  Albone, Earl
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C8
;  CURRENT APPLICATION NUMBER: US/09/907,969
;  CURRENT FILING DATE:  2001-07-17
;  NUMBER OF SEQ ID NOS: 596
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-907-969-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 163
US-09-827-271-100
; Sequence 100, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 164
US-10-198-053-100
; Sequence 100, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-100

  Query Match             8.7%;  Score 191;  DB 6;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 165
US-10-860-790-100
; Sequence 100, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
```

```
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C11
;  CURRENT APPLICATION NUMBER: US/10/860,790
;  CURRENT FILING DATE:  2004-06-02
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-860-790-100

  Query Match             8.7%;  Score 191;  DB 8;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-92;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 166
US-10-242-535A-24536
; Sequence 24536, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24536
;   LENGTH: 355
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (108)..(108)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (286)..(286)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (288)..(288)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-24536

  Query Match             8.3%;  Score 183;  DB 7;  Length 355;
  Best Local Similarity   99.3%;  Pred. No. 1e-87;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 60

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954
             ||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 GCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGNGGAGCCCAGTGA 120

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 180

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 240

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 285


RESULT 167
US-10-085-783A-24536
; Sequence 24536, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24536
;   LENGTH: 355
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (108)..(108)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (286)..(286)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (288)..(288)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-24536

  Query Match             8.3%;  Score 183;  DB 7;  Length 355;
```

```
  Best Local Similarity   99.3%;  Pred. No. 1e-87;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 60

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954
             ||| |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 GCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGNGGAGCCCAGTGA 120

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 180

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 240

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 2119
             ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGG 285


RESULT 168
US-10-242-535A-5700
; Sequence 5700, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5700
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-5700

  Query Match             8.3%;  Score 182;  DB 7;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-87;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1853 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 140

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         141 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 200

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 260

Qy       2033 GA 2034
             ||
Db         261 GA 262


RESULT 169
US-10-085-783A-5700
; Sequence 5700, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5700
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-5700

  Query Match             8.3%;  Score 182;  DB 7;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-87;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1853 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 140

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         141 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 200

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 260

Qy       2033 GA 2034
             ||
Db         261 GA 262


RESULT 170
US-09-777-564-1048
; Sequence 1048, Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.493
;  CURRENT APPLICATION NUMBER: US/09/777,564
;  CURRENT FILING DATE:  2001-02-05
;  NUMBER OF SEQ ID NOS:  1730
;  SOFTWARE:  FastSEQ for Window Version 4.0
; SEQ ID NO 1048
;   LENGTH: 227
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-777-564-1048

  Query Match             8.0%;  Score 176;  DB 3;  Length 227;
  Best Local Similarity   99.6%;  Pred. No. 6.1e-84;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 60
```

```
Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 TGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 120

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 180

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 227
```

RESULT 171
US-10-015-219-1048
; Sequence 1048, Application US/10015219
; Publication No. US20030008299A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493C1
; CURRENT APPLICATION NUMBER: US/10/015,219
; CURRENT FILING DATE: 2002-03-02
; NUMBER OF SEQ ID NOS: 1739
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1048
; LENGTH: 227
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-219-1048

Query Match 8.0%; Score 176; DB 5; Length 227;
Best Local Similarity 99.6%; Pred. No. 6.1e-84;
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1813 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 60

Qy       1873 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1932
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 TGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 120

Qy       1933 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 180

Qy       1993 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 2039
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 227
```

RESULT 172
US-10-357-930-16069/c
; Sequence 16069, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel. Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16069
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 263, 290
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-16069

Query Match 7.8%; Score 171; DB 8; Length 294;
Best Local Similarity 100.0%; Pred. No. 3e-81;
Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        518 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 203

Qy        578 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCC 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCC 143

Qy        638 TGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 92
```

RESULT 173
US-10-242-535A-36526
; Sequence 36526, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36526
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (170)..(170)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (184)..(184)

```
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (253)..(253)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (265)..(265)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-36526

  Query Match             7.3%;  Score 161;  DB 7;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-76;
  Matches  161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         34 TCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGA 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 TCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGA 68

Qy         94 CTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 128

Qy        154 AGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              |||||||||||||||||||||||||||||||||||||||||
Db        129 AGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 169


RESULT 174
US-10-085-783A-36526
; Sequence 36526, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36526
;   LENGTH: 273
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (170)..(170)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (184)..(184)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (253)..(253)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (265)..(265)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-36526

  Query Match             7.3%;  Score 161;  DB 7;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-76;
  Matches  161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         34 TCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGA 93
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 TCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGA 68

Qy         94 CTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 128

Qy        154 AGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              |||||||||||||||||||||||||||||||||||||||||
Db        129 AGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 169


RESULT 175
US-10-242-535A-50471
; Sequence 50471, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50471
;   LENGTH: 211
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (14)..(14)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (27)..(27)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)..(51)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-50471

  Query Match             7.3%;  Score 160;  DB 7;  Length 211;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-75;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 111

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 171

Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
                ||||||||||||||||||||||||||||||||||||||||
Db        172 TTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 211


RESULT 176
US-10-085-783A-50471
; Sequence 50471, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50471
;   LENGTH: 211
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (14)..(14)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (27)..(27)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (51)..(51)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-50471

  Query Match             7.3%;  Score 160;  DB 7;  Length 211;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-75;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 111

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 171

Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
                ||||||||||||||||||||||||||||||||||||||||
Db        172 TTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 211


RESULT 177
US-09-969-034-3133/c
; Sequence 3133, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
;  APPLICANT: Burgess, Christopher C.
;  APPLICANT:  Astle, Jon H.
;  APPLICANT:  Carroll, Eddie III
;  APPLICANT:  Catino, Theodore J.
;  APPLICANT:  Dwivedi, Poornima
;  APPLICANT:  Molino, Gary A.
;  APPLICANT:  Thiagalingam, Arunthathi
;  APPLICANT:  Lewis, Marcia E.
;  TITLE OF INVENTION: Nucleic Acid Sequences Differentially
;  TITLE OF INVENTION:  Expressed in Cancer Tissue
;  FILE REFERENCE: 1657/1032
;  CURRENT APPLICATION NUMBER: US/09/969,034
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: 60/237,271
;  PRIOR FILING DATE: 2000-02-10
;  NUMBER OF SEQ ID NOS: 4494
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3133
;   LENGTH: 509
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 262, 308, 328, 339, 348, 371, 389, 395, 404, 407, 417, 425,
;   LOCATION: 486, 497, 503
;   OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3133

  Query Match             6.9%;  Score 152;  DB 3;  Length 509;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-71;
  Matches  152;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2041 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGG 115

Qy       2101 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 TCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTG 55

Qy       2161 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
                ||||||||||||||||||||||||||||||||
Db         54 CACTTTCCTTTCAATAAAGTTGTTGCATTCCC 23


RESULT 178
US-10-242-535A-8748
; Sequence 8748, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8748
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-8748
```

```
  Query Match             6.5%;  Score 143;  DB 7;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-66;
  Matches  143;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 176

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 236

Qy       2019 TTGATCTTTGCTGGGAAACAGCT 2041
             |||||||||||||||||||||||
Db        237 TTGATCTTTGCTGGGAAACAGCT 259


RESULT 179
US-10-085-783A-8748
; Sequence 8748, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8748
;   LENGTH: 400
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-8748

  Query Match             6.5%;  Score 143;  DB 7;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-66;
  Matches  143;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 176

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 236

Qy       2019 TTGATCTTTGCTGGGAAACAGCT 2041
             |||||||||||||||||||||||
Db        237 TTGATCTTTGCTGGGAAACAGCT 259


RESULT 180
US-10-264-049-1008
; Sequence 1008, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1008
;   LENGTH: 794
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (174)..(174)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (214)..(214)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (217)..(217)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (234)..(234)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (255)..(256)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (284)..(284)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (315)..(315)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (333)..(333)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (389)..(390)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (416)..(416)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (418)..(418)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (423)..(423)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (426)..(426)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (434)..(434)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (452)..(452)
;   OTHER INFORMATION: n equals a,t,g, or c
;   FEATURE:
;   NAME/KEY: misc_feature
```

; LOCATION: (457)..(457)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (462)..(462)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (466)..(466)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (468)..(468)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (488)..(488)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (491)..(491)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (510)..(510)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (515)..(515)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (545)..(545)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (547)..(547)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (553)..(553)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (556)..(557)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (573)..(574)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (585)..(585)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (591)..(591)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (599)..(599)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (602)..(602)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (607)..(607)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (611)..(611)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (662)..(662)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (667)..(667)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (684)..(684)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (717)..(717)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (731)..(731)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (736)..(736)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (740)..(741)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (755)..(755)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (759)..(759)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature

Query Match 6.4%; Score 141; DB 6; Length 794;
Best Local Similarity 100.0%; Pred. No. 4.5e-65;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1442 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 1501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACAC 60

Qy       1502 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 120

Qy       1562 GTTGATCTTTGCTGGGAAACA 1582
             |||||||||||||||||||||
Db        121 GTTGATCTTTGCTGGGAAACA 141
```

RESULT 181
US-09-777-564-1207/c
; Sequence 1207, Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05

```
;  NUMBER OF SEQ ID NOS:  1730
;  SOFTWARE:  FastSEQ for Window Version 4.0
; SEQ ID NO 1207
;   LENGTH: 240
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(240)
;   OTHER INFORMATION: n = A,T,C or G
US-09-777-564-1207

  Query Match             6.2%;  Score 137;  DB 3;  Length 240;
  Best Local Similarity   99.5%;  Pred. No. 6.6e-63;
  Matches  187;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 129

Qy       1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        128 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 69

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 9

Qy       1575 GGGAAACA 1582
             ||||||||
Db           8 GGGAAACA 1


RESULT 182
US-10-015-219-1207/c
; Sequence 1207, Application US/10015219
; Publication No. US20030008299A1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.493C1
;  CURRENT APPLICATION NUMBER: US/10/015,219
;  CURRENT FILING DATE:  2002-03-02
;  NUMBER OF SEQ ID NOS: 1739
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1207
;   LENGTH: 240
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 189, 201, 230, 232, 233
;   OTHER INFORMATION: n = A,T,C or G
US-10-015-219-1207

  Query Match             6.2%;  Score 137;  DB 5;  Length 240;
  Best Local Similarity   99.5%;  Pred. No. 6.6e-63;
  Matches  187;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1395 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 129

Qy       1455 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1514
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        128 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 69

Qy       1515 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 9

Qy       1575 GGGAAACA 1582
             ||||||||
Db           8 GGGAAACA 1


RESULT 183
US-10-242-535A-45031
; Sequence 45031, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45031
;   LENGTH: 208
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-45031

  Query Match             6.2%;  Score 135;  DB 7;  Length 208;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-62;
  Matches  135;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1491 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 60

Qy       1551 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 120

Qy       1611 TACAACATCCAGAAA 1625
             |||||||||||||||
Db        121 TACAACATCCAGAAA 135


RESULT 184
US-10-085-783A-45031
; Sequence 45031, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45031
;   LENGTH: 208
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-45031
```

```
  Query Match             6.2%;  Score 135;  DB 7;  Length 208;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-62;
  Matches  135;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1491 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 60

Qy       1551 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 120

Qy       1611 TACAACATCCAGAAA 1625
             |||||||||||||||
Db        121 TACAACATCCAGAAA 135


RESULT 185
US-10-913-447-67
; Sequence 67, Application US/10913447
; Publication No. US20050118615A1
; GENERAL INFORMATION:
;  APPLICANT: Anat OREN et al.
;  TITLE OF INVENTION: CHORDIN LIKE HOMOLOGS
;  FILE REFERENCE: 2786-0247PUS1
;  CURRENT APPLICATION NUMBER: US/10/913,447
;  CURRENT FILING DATE:  2004-08-09
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
;   LENGTH: 133
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Ubiquitin Amplicon
US-10-913-447-67

  Query Match             6.1%;  Score 133;  DB 9;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-61;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCA 152
             |||||||||||||
Db        121 GAATGTCAAGGCA 133


RESULT 186
US-11-034-010-33
; Sequence 33, Application US/11034010
; Publication No. US20050186600A1
; GENERAL INFORMATION:
;  APPLICANT: SELLA-TAVOR, Osnat
;  APPLICANT:  ROTMAN, Galit
;  APPLICANT:  POLLOCK, Sarah
;  APPLICANT:  DIBER, Alex
;  APPLICANT:  WALACH, Shira
;  APPLICANT:  SAMEACH-GREENWALD, Shirley
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL UBCH10 POLYPEPTIDES AND KITS AND
;  TITLE OF INVENTION:  METHODS USING SAME
;  FILE REFERENCE: 28482
;  CURRENT APPLICATION NUMBER: US/11/034,010
;  CURRENT FILING DATE:  2005-01-13
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33
;   LENGTH: 133
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Ubiquitin amplicon
US-11-034-010-33

  Query Match             6.1%;  Score 133;  DB 10;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 9.5e-61;
  Matches  133;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCA 152
             |||||||||||||
Db        121 GAATGTCAAGGCA 133


RESULT 187
US-10-242-535A-34635
; Sequence 34635, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34635
;   LENGTH: 374
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (33)..(33)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (166)..(166)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (185)..(185)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (200)..(200)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (227)..(227)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (260)..(260)
;   OTHER INFORMATION: n is a, c, g, or t
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (290)..(290)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (296)..(296)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (311)..(311)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (368)..(369)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-34635

  Query Match             6.0%;  Score 132;  DB 7;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-60;
  Matches  132;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        178 CTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGT 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 CTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGT 93

Qy        238 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 153

Qy        298 TGCAAATCTTCG 309
              ||||||||||||
Db        154 TGCAAATCTTCG 165


RESULT 188
US-10-085-783A-34635
; Sequence 34635, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34635
;   LENGTH: 374
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (33)..(33)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (166)..(166)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (185)..(185)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (200)..(200)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (227)..(227)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (260)..(260)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (290)..(290)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (296)..(296)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (311)..(311)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (368)..(369)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-34635

  Query Match             6.0%;  Score 132;  DB 7;  Length 374;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-60;
  Matches  132;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        178 CTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGT 237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 CTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGT 93

Qy        238 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 153

Qy        298 TGCAAATCTTCG 309
              ||||||||||||
Db        154 TGCAAATCTTCG 165


RESULT 189
US-10-242-535A-4391
; Sequence 4391, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 4391
;   LENGTH: 168
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-4391

  Query Match             5.7%;  Score 125;  DB 7;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 2e-56;
  Matches  125;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1423 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 1482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 60

Qy       1483 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 1542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 120

Qy       1543 TCCCT 1547
             |||||
Db        121 TCCCT 125


RESULT 190
US-10-085-783A-4391
; Sequence 4391, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4391
;   LENGTH: 168
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-4391

  Query Match             5.7%;  Score 125;  DB 7;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 2e-56;
  Matches  125;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1423 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 1482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCG 60

Qy       1483 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 1542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCA 120

Qy       1543 TCCCT 1547
             |||||
Db        121 TCCCT 125


RESULT 191
US-09-884-441-99/c
; Sequence 99, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C7
;  CURRENT APPLICATION NUMBER: US/09/884,441
;  CURRENT FILING DATE:  2001-06-18
;  NUMBER OF SEQ ID NOS: 489
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-884-441-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-55;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 192
US-09-907-969-99/c
; Sequence 99, Application US/09907969
; Publication No. US20030091580A1
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Hill, Paul
;  APPLICANT:  Albone, Earl
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C8
;  CURRENT APPLICATION NUMBER: US/09/907,969
;  CURRENT FILING DATE:  2001-07-17
;  NUMBER OF SEQ ID NOS: 596
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-907-969-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-55;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457
```

RESULT 193
US-09-827-271-99/c
; Sequence 99, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-827-271-99

Query Match 5.6%; Score 123; DB 3; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.3e-55;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
Db 579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy 1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
Db 519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy 1508 GAA 1510
Db 459 GAA 457

RESULT 194
US-10-198-053-99/c
; Sequence 99, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-053-99

Query Match 5.6%; Score 123; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.3e-55;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
Db 579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy 1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
Db 519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy 1508 GAA 1510
Db 459 GAA 457

RESULT 195
US-10-860-790-99/c
; Sequence 99, Application US/10860790
; Publication No. US20050031634A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C11
; CURRENT APPLICATION NUMBER: US/10/860,790
; CURRENT FILING DATE: 2004-06-02
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-860-790-99

Query Match 5.6%; Score 123; DB 8; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.3e-55;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
Db 579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy 1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
Db 519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy 1508 GAA 1510
Db 459 GAA 457

RESULT 196
US-10-242-535A-34587
; Sequence 34587, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34587
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (172)..(172)
; OTHER INFORMATION: n is a, c, g, or t

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (217)..(217)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (274)..(274)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (298)..(298)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (304)..(304)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (313)..(313)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-34587

  Query Match             5.5%;  Score 120;  DB 7;  Length 341;
  Best Local Similarity   99.4%;  Pred. No. 9.7e-54;
  Matches  170;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        218 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 60

Qy        278 GCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 337
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         61 GCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCGCTGGCAAGACCATCAC 120

Qy        338 CCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 171


RESULT 197
US-10-085-783A-34587
; Sequence 34587, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34587
;   LENGTH: 341
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (172)..(172)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (217)..(217)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (274)..(274)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (298)..(298)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (304)..(304)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (313)..(313)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-34587

  Query Match             5.5%;  Score 120;  DB 7;  Length 341;
  Best Local Similarity   99.4%;  Pred. No. 9.7e-54;
  Matches  170;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        218 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 60

Qy        278 GCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 337
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         61 GCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCGCTGGCAAGACCATCAC 120

Qy        338 CCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 171


RESULT 198
US-09-918-995-28329
; Sequence 28329, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28329
;   LENGTH: 469
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(469)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28329

  Query Match             5.4%;  Score 118;  DB 3;  Length 469;
  Best Local Similarity   99.4%;  Pred. No. 1.2e-52;
  Matches  168;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2024 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 2083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 245

Qy       2084 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 2143
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        246 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAATTTCCCCTTTTAAGGT 305

Qy       2144 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 354
```

```
RESULT 199
US-10-242-535A-28125
; Sequence 28125, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28125
;   LENGTH: 112
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-28125

  Query Match             5.1%;  Score 112;  DB 7;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 2e-49;
  Matches  112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 60

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 112


RESULT 200
US-10-085-783A-28125
; Sequence 28125, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28125
;   LENGTH: 112
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-28125

  Query Match             5.1%;  Score 112;  DB 7;  Length 112;
  Best Local Similarity   100.0%;  Pred. No. 2e-49;
  Matches  112;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 60

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 112


RESULT 201
US-10-242-535A-4938
; Sequence 4938, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4938
;   LENGTH: 167
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-4938

  Query Match             5.0%;  Score 109;  DB 7;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-48;
  Matches  109;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2084 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 2143
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 60

Qy       2144 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 109


RESULT 202
US-10-085-783A-4938
; Sequence 4938, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4938
;   LENGTH: 167
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-4938

  Query Match             5.0%;  Score 109;  DB 7;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-48;
```

```
  Matches  109;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2084 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGT 60

Qy       2144 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 109


RESULT 203
US-10-393-602-140/c
; Sequence 140, Application US/10393602
; Publication No. US20030170714A1
;   GENERAL INFORMATION:
;        APPLICANT: Gregory Dolganov
;        TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
;                            Polypeptides
;        NUMBER OF SEQUENCES: 151
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Dehlinger & Associates
;             STREET: 350 Cambridge Avenue, Suite 250
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94306
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/393,602
;             FILING DATE: 19-Mar-2003
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/592,126
;             FILING DATE: 26-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Sholtz, Charles K.
;             REGISTRATION NUMBER: 38,615
;             REFERENCE/DOCKET NUMBER: 4600-0111
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 324-0880
;             TELEFAX: (415) 324-0960
;   INFORMATION FOR SEQ ID NO: 140:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 795 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA to mRNA
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             INDIVIDUAL ISOLATE: E67.seq
;        SEQUENCE DESCRIPTION: SEQ ID NO: 140:
US-10-393-602-140

  Query Match             4.7%;  Score 104;  DB 6;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-45;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             |||||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6


RESULT 204
US-10-242-535A-13237
; Sequence 13237, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13237
;   LENGTH: 107
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-13237

  Query Match             4.6%;  Score 100;  DB 7;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 6e-43;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 2149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 67

Qy       2150 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATT 2189
             |||||||||||||||||||||||||||||||||||||||||
Db         68 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATT 107


RESULT 205
US-10-085-783A-13237
; Sequence 13237, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13237
;   LENGTH: 107
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-13237

  Query Match             4.6%;  Score 100;  DB 7;  Length 107;
  Best Local Similarity   100.0%;  Pred. No. 6e-43;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2090 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 2149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAAC 67
```

```
Qy       2150 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATT 2189
              ||||||||||||||||||||||||||||||||||||||||
Db         68 AAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATT 107
```

RESULT 206
US-09-969-034-3083
; Sequence 3083, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3083
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12, 23, 31, 297, 301, 307, 313, 331, 341, 352, 357, 394,
; LOCATION: 400, 404, 439, 445, 455, 456, 473, 474, 476, 479, 490, 498,
; LOCATION: 500, 511
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3083

Query Match 4.6%; Score 100; DB 3; Length 520;
Best Local Similarity 100.0%; Pred. No. 5.9e-43;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         32 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 91

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
              |||||||||||||||||||||||||||||||||||||||
Db         92 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 131
```

Search completed: March 7, 2006, 00:44:40
Job time : 1799 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 21:06:42 ; Search time 657 Seconds
(without alignments)
7314.958 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 7218535 seqs, 1096242582 residues

Word size : 100

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2177	99.3	2201	9	US-11-050-857-1056	Sequence 1056, Ap
2	2177	99.3	2201	12	US-11-043-590-46	Sequence 46, Appl
3	433	19.8	460	12	US-11-108-172-163	Sequence 163, App
4	303	13.8	412	12	US-11-108-172-660	Sequence 660, App
5	133	6.1	133	9	US-11-050-857-1059	Sequence 1059, Ap
6	133	6.1	133	12	US-11-043-590-38	Sequence 38, Appl
7	133	6.1	133	12	US-11-043-788-445	Sequence 445, App

ALIGNMENTS

RESULT 1
US-11-050-857-1056
; Sequence 1056, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1056
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-050-857-1056

Query Match 99.3%; Score 2177; DB 9; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900
```

```
Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy       1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy       1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy       1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy       1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy       1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy       1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy       1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy       1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy       2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy       2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy       2176 AAAGTTGTTGCATTCCC 2192
              |||||||||||||||||
Db       2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 2
US-11-043-590-46
; Sequence 46, Application US/11043590
; Publication No. US20050277156A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use there
; FILE REFERENCE: 1847.1011
; CURRENT APPLICATION NUMBER: US/11/043,590
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 46
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-043-590-46

Query Match 99.3%; Score 2177; DB 12; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy          76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy         136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy         196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy         376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy         436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy         496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540
```

```
Qy       556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
Db       541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy       616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
Db       601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy       676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
Db       661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy       736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
Db       721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy       796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
Db       781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy       856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
Db       841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy       916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
Db       901 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 960

Qy       976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
Db       961 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1020

Qy      1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
Db      1021 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1080

Qy      1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
Db      1081 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1140

Qy      1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
Db      1141 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1200

Qy      1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
Db      1201 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1260

Qy      1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
Db      1261 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1320

Qy      1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
Db      1321 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1380

Qy      1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
Db      1381 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1440

Qy      1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
Db      1441 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1500

Qy      1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
Db      1501 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1560

Qy      1576 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
Db      1561 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1620

Qy      1636 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1695
Db      1621 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1680

Qy      1696 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1755
Db      1681 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1740

Qy      1756 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1815
Db      1741 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGG 1800

Qy      1816 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1875
Db      1801 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1860

Qy      1876 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
Db      1861 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1920

Qy      1936 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1995
Db      1921 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 1980

Qy      1996 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2055
Db      1981 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 2040

Qy      2056 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2115
Db      2041 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGG 2100

Qy      2116 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2175
Db      2101 GGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAAT 2160

Qy      2176 AAAGTTGTTGCATTCCC 2192
Db      2161 AAAGTTGTTGCATTCCC 2177
```

RESULT 3
US-11-108-172-163
; Sequence 163, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10

```
;  PRIOR APPLICATION NUMBER: US 09/649,811
;  PRIOR FILING DATE: 2000-08-28
;  PRIOR APPLICATION NUMBER: US 09/609,448
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: US 09/575,251
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: US 09/519,444
;  PRIOR FILING DATE: 2000-03-06
;  PRIOR APPLICATION NUMBER: US 09/504,629
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: US 09/480,321
;  PRIOR FILING DATE: 2000-01-10
;  PRIOR APPLICATION NUMBER: US 09/476,296
;  PRIOR FILING DATE: 1999-12-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1130
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 163
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-108-172-163

  Query Match             19.8%;  Score 433;  DB 12;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-220;
  Matches  433;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
              |||||||||||||
Db        448 GCTGGAAGATGGG 460


RESULT 4
US-11-108-172-660
; Sequence 660, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Smith, Carole L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Clapper, Jonathan D.
;  APPLICANT:  Skeiky, Yasir A. W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;  TITLE OF INVENTION:  OF COLON CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.471C15
;  CURRENT APPLICATION NUMBER: US/11/108,172
;  CURRENT FILING DATE:  2005-04-15
;  PRIOR APPLICATION NUMBER: US 10/025,380
;  PRIOR FILING DATE: 2001-12-19
;  PRIOR APPLICATION NUMBER: US 09/922,217
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 09/833,263
;  PRIOR FILING DATE: 2001-04-10
;  PRIOR APPLICATION NUMBER: US 09/649,811
;  PRIOR FILING DATE: 2000-08-28
;  PRIOR APPLICATION NUMBER: US 09/609,448
;  PRIOR FILING DATE: 2000-06-29
;  PRIOR APPLICATION NUMBER: US 09/575,251
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: US 09/519,444
;  PRIOR FILING DATE: 2000-03-06
;  PRIOR APPLICATION NUMBER: US 09/504,629
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: US 09/480,321
;  PRIOR FILING DATE: 2000-01-10
;  PRIOR APPLICATION NUMBER: US 09/476,296
;  PRIOR FILING DATE: 1999-12-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1130
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 660
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-108-172-660

  Query Match             13.8%;  Score 303;  DB 12;  Length 412;
  Best Local Similarity   99.7%;  Pred. No. 2.9e-151;
  Matches  353;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 118

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 178

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 238

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 298

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        299 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 358

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 412


RESULT 5
US-11-050-857-1059
```

; Sequence 1059, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 1847.1005
; CURRENT APPLICATION NUMBER: US/11/050,857
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1059
; LENGTH: 133
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
US-11-050-857-1059

Query Match 6.1%; Score 133; DB 9; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e-60;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy 80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy 140 GAATGTCAAGGCA 152
||||||||||||||
Db 121 GAATGTCAAGGCA 133

RESULT 6
US-11-043-590-38
; Sequence 38, Application US/11043590
; Publication No. US20050277156A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use there
; FILE REFERENCE: 1847.1011
; CURRENT APPLICATION NUMBER: US/11/043,590
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 38
; LENGTH: 133
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide
US-11-043-590-38

Query Match 6.1%; Score 133; DB 12; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e-60;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy 80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy 140 GAATGTCAAGGCA 152
||||||||||||||
Db 121 GAATGTCAAGGCA 133

RESULT 7
US-11-043-788-445

; Sequence 445, Application US/11043788
; Publication No. US20060014166A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHO
; TITLE OF INVENTION: THEREOF FOR DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: 1847.1006
; CURRENT APPLICATION NUMBER: US/11/043,788
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 506
; SEQ ID NO 445
; LENGTH: 133
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Ubiquitin Amplicon
US-11-043-788-445

Query Match 6.1%; Score 133; DB 12; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.4e-60;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy 80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy 140 GAATGTCAAGGCA 152
||||||||||||||
Db 121 GAATGTCAAGGCA 133

Search completed: March 6, 2006, 21:38:00
Job time : 658 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 20:51:30 ; Search time 411 Seconds
(without alignments)
9480.329 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 100

Total number of hits satisfying chosen parameters: 51

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	449	20.5	547	3	US-09-370-838-171	Sequence 171, App
2	449	20.5	547	3	US-09-854-133-171	Sequence 171, App
3	433	19.8	460	3	US-09-401-064-163	Sequence 163, App
4	430	19.6	494	3	US-09-370-838-275	Sequence 275, App
5	430	19.6	494	3	US-09-854-133-275	Sequence 275, App
6	426	19.4	483	3	US-09-370-838-141	Sequence 141, App
7	426	19.4	483	3	US-09-854-133-141	Sequence 141, App
8	411	18.8	412	3	US-09-854-133-692	Sequence 692, App
9	410	18.7	434	3	US-09-370-838-248	Sequence 248, App
10	410	18.7	434	3	US-09-854-133-248	Sequence 248, App
11	370	16.9	518	3	US-09-404-879A-152	Sequence 152, App
12	370	16.9	518	3	US-09-338-933-152	Sequence 152, App
13	370	16.9	518	3	US-09-215-681-152	Sequence 152, App
14	370	16.9	518	3	US-09-216-003A-152	Sequence 152, App
15	370	16.9	518	3	US-09-667-857-152	Sequence 152, App
16	370	16.9	518	3	US-10-198-053-152	Sequence 152, App
17	370	16.9	518	3	US-09-827-271-152	Sequence 152, App
18	270	12.3	284	3	US-09-506-729-52	Sequence 52, Appl
19	248	11.3	585	3	US-09-404-879A-45	Sequence 45, Appl
20	248	11.3	585	3	US-09-338-933-45	Sequence 45, Appl
21	248	11.3	585	3	US-09-215-681-45	Sequence 45, Appl
22	248	11.3	585	3	US-09-216-003A-45	Sequence 45, Appl
23	248	11.3	585	3	US-09-667-857-45	Sequence 45, Appl
24	248	11.3	585	3	US-10-198-053-45	Sequence 45, Appl
25	248	11.3	585	3	US-09-827-271-45	Sequence 45, Appl
26	211	9.6	440	3	US-09-397-787-277	Sequence 277, App
27	207	9.4	228	3	US-10-237-551-80	Sequence 80, Appl
c 28	207	9.4	609	3	US-09-385-982-169	Sequence 169, App
c 29	191	8.7	605	3	US-09-404-879A-95	Sequence 95, Appl
c 30	191	8.7	605	3	US-09-338-933-95	Sequence 95, Appl
c 31	191	8.7	605	3	US-09-215-681-95	Sequence 95, Appl
c 32	191	8.7	605	3	US-09-216-003A-95	Sequence 95, Appl
c 33	191	8.7	605	3	US-09-667-857-95	Sequence 95, Appl
c 34	191	8.7	605	3	US-10-198-053-95	Sequence 95, Appl
c 35	191	8.7	605	3	US-09-827-271-95	Sequence 95, Appl
36	191	8.7	697	3	US-09-404-879A-100	Sequence 100, App
37	191	8.7	697	3	US-09-338-933-100	Sequence 100, App
38	191	8.7	697	3	US-09-215-681-100	Sequence 100, App
39	191	8.7	697	3	US-09-216-003A-100	Sequence 100, App
40	191	8.7	697	3	US-09-667-857-100	Sequence 100, App
41	191	8.7	697	3	US-10-198-053-100	Sequence 100, App
42	191	8.7	697	3	US-09-827-271-100	Sequence 100, App
c 43	123	5.6	636	3	US-09-404-879A-99	Sequence 99, Appl
c 44	123	5.6	636	3	US-09-338-933-99	Sequence 99, Appl
c 45	123	5.6	636	3	US-09-215-681-99	Sequence 99, Appl
c 46	123	5.6	636	3	US-09-216-003A-99	Sequence 99, Appl
c 47	123	5.6	636	3	US-09-667-857-99	Sequence 99, Appl
c 48	123	5.6	636	3	US-10-198-053-99	Sequence 99, Appl
c 49	123	5.6	636	3	US-09-827-271-99	Sequence 99, Appl
c 50	104	4.7	795	2	US-08-592-126-140	Sequence 140, App
c 51	104	4.7	795	3	US-09-168-595-140	Sequence 140, App

ALIGNMENTS

RESULT 1
US-09-370-838-171
; Sequence 171, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-171

Query Match 20.5%; Score 449; DB 3; Length 547;
Best Local Similarity 99.8%; Pred. No. 1.3e-212;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
```

```
Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348

Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408

Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468

Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528

Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548

Db        528 AGATCTTCGTGAAGACCCTG 547
```

RESULT 2
US-09-854-133-171
; Sequence 171, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-171

Query Match 20.5%; Score 449; DB 3; Length 547;
Best Local Similarity 99.8%; Pred. No. 1.3e-212;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108

Db         48 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 107

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168

Db        108 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 167

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228

Db        168 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 227

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288

Db        228 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 287

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348

Db        288 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 347

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408

Db        348 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 407

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468

Db        408 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 467

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528

Db        468 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 527

Qy        529 AGATCTTCGTGAAGACCCTG 548

Db        528 AGATCTTCGTGAAGACCCTG 547
```

RESULT 3
US-09-401-064-163
; Sequence 163, Application US/09401064
; Patent No. 6623923
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C2
; CURRENT APPLICATION NUMBER: US/09/401,064
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 163
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-401-064-163

Query Match 19.8%; Score 433; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.1e-204;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82

Db         28 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 87

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142

Db         88 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 147

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202

Db        148 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 207

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262

Db        208 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 267

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322

Db        268 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 327

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382

Db        328 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 387

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442

Db        388 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 447

Qy        443 GCTGGAAGATGGG 455
```

```
                ||||||||||||||
Db        448 GCTGGAAGATGGG 460
```

RESULT 4
US-09-370-838-275
; Sequence 275, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 275
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (379)
; OTHER INFORMATION: n=A,T,C or G
US-09-370-838-275

Query Match 19.6%; Score 430; DB 3; Length 494;
Best Local Similarity 99.8%; Pred. No. 3.5e-203;
Matches 480; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494
```

RESULT 5
US-09-854-133-275
; Sequence 275, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 275
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (379)
; OTHER INFORMATION: n=A,T,C or G
US-09-854-133-275

Query Match 19.6%; Score 430; DB 3; Length 494;
Best Local Similarity 99.8%; Pred. No. 3.5e-203;
Matches 480; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 73

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 133

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 193

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 253

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 313

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 373

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 ATCCANGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 433

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 493

Qy        504 G 504
              |
Db        494 G 494
```

RESULT 6
US-09-370-838-141
; Sequence 141, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:

```
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Lodes, Michael J.
;  APPLICANT: Mohamath, Roadoh
;  APPLICANT: Secrist, Heather
;  TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
;  TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.475C1
;  CURRENT APPLICATION NUMBER: US/09/370,838
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: US 09/285,323
;  EARLIER FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 289
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 141
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-370-838-141

  Query Match             19.4%;  Score 426;  DB 3;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 7
US-09-854-133-141
; Sequence 141, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 141
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-854-133-141

  Query Match             19.4%;  Score 426;  DB 3;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-201;
  Matches  426;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         58 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 117

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        118 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 177

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        178 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 237

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 297

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 357

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 417

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        418 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 477

Qy        649 GGTTGA 654
              ||||||
Db        478 GGTTGA 483


RESULT 8
US-09-854-133-692
; Sequence 692, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 692
;   LENGTH: 412
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-854-133-692

  Query Match             18.8%;  Score 411;  DB 3;  Length 412;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-194;
  Matches  411;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         35 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGAC 61
```

```
Qy          95 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 121

Qy         155 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 GATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACA 181

Qy         215 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 241

Qy         275 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 GGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 301

Qy         335 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 CACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAA 361

Qy         395 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 GGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 412
```

RESULT 9
US-09-370-838-248
; Sequence 248, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 248
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-370-838-248

Query Match 18.7%; Score 410; DB 3; Length 434;
Best Local Similarity 100.0%; Pred. No. 2.9e-193;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy          92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy         152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy         212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy         272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy         332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy         392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 10
US-09-854-133-248
; Sequence 248, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 248
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-248

Query Match 18.7%; Score 410; DB 3; Length 434;
Best Local Similarity 100.0%; Pred. No. 2.9e-193;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 60

Qy          92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 120

Qy         152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 180

Qy         212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 240

Qy         272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 300

Qy         332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 360

Qy         392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 410
```

RESULT 11
US-09-404-879A-152
; Sequence 152, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.

```
;  APPLICANT:  Algate, Paul A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462C2
;  CURRENT APPLICATION NUMBER: US/09/404,879A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  393
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-404-879A-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 12
US-09-338-933-152
; Sequence 152, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer Lynn
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: OVARIAN CANCER
;  FILE REFERENCE: 210121.462C1
;  CURRENT APPLICATION NUMBER: US/09/338,933
;  CURRENT FILING DATE: 1999-06-23
;  NUMBER OF SEQ ID NOS: 312
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-338-933-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 13
US-09-215-681-152
; Sequence 152, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: Frudakis, Tony N.
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
;  TITLE OF INVENTION: OF OVARIAN CANCER
;  FILE REFERENCE: 210121.463
;  CURRENT APPLICATION NUMBER: US/09/215,681A
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 310
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-215-681-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
```

```
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018

Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078

Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138

Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190

Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 14
US-09-216-003A-152
; Sequence 152, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 152
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-152

Query Match 16.9%; Score 370; DB 3; Length 518;
Best Local Similarity 99.6%; Pred. No. 1.9e-173;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778

Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838

Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898

Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958

Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018

Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078

Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138

Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190

Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 15
US-09-667-857-152
; Sequence 152, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 152
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-667-857-152

Query Match 16.9%; Score 370; DB 3; Length 518;
Best Local Similarity 99.6%; Pred. No. 1.9e-173;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778

Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838

Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898

Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958

Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018

Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078

Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138

Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190

Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518
```

RESULT 16
US-10-198-053-152
; Sequence 152, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:

```
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 17
US-09-827-271-152
; Sequence 152, Application US/09827271
; Patent No. 6962980
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 152
;   LENGTH: 518
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-152

  Query Match             16.9%;  Score 370;  DB 3;  Length 518;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-173;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         47 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 106

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
             |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 GACCAGCAKAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 166

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 226

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 286

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 346

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 406

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 466

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 2190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTC 518


RESULT 18
US-09-506-729-52
; Sequence 52, Application US/09506729
; Patent No. 6365352
; GENERAL INFORMATION:
;  APPLICANT: Yerramilli, Subrahmanyam V.
;  APPLICANT: Prashar, Yatindra
;  APPLICANT: Newberger, Peter
;  APPLICANT: Goguen, Jon
;  APPLICANT: Weissman, Sherman M.
;  TITLE OF INVENTION: A PROCESS TO STUDY CHANGES IN GENE EXPRESSION IN
;  TITLE OF INVENTION: GRANULOCYTIC CELLS
;  FILE REFERENCE: 44921-5016-US
;  CURRENT APPLICATION NUMBER: US/09/506,729
;  CURRENT FILING DATE: 2000-02-18
;  EARLIER APPLICATION NUMBER: PCT/US98/17284
;  EARLIER FILING DATE: 1998-08-21
;  EARLIER APPLICATION NUMBER: 60/056,844
;  EARLIER FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-506-729-52

  Query Match             12.3%;  Score 270;  DB 3;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-124;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 74
```

```
Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 134

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 194

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 2152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAA 254

Qy       2153 TTTCATTGCACTTTCCTTTCAATAAAGTTG 2182
             ||||||||||||||||||||||||||||||
Db        255 TTTCATTGCACTTTCCTTTCAATAAAGTTG 284


RESULT 19
US-09-404-879A-45
; Sequence 45, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462C2
;  CURRENT APPLICATION NUMBER: US/09/404,879A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  393
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-404-879A-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db         532 AAGGTTTC 539


RESULT 20
US-09-338-933-45
; Sequence 45, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer Lynn
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: OVARIAN CANCER
;  FILE REFERENCE: 210121.462C1
;  CURRENT APPLICATION NUMBER: US/09/338,933
;  CURRENT FILING DATE: 1999-06-23
;  NUMBER OF SEQ ID NOS: 312
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-338-933-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db         532 AAGGTTTC 539


RESULT 21
US-09-215-681-45
; Sequence 45, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: Frudakis, Tony N.
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
;  TITLE OF INVENTION: OF OVARIAN CANCER
;  FILE REFERENCE: 210121.463
;  CURRENT APPLICATION NUMBER: US/09/215,681A
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 310
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-215-681-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
                ||||||||
Db        532 AAGGTTTC 539


RESULT 22
US-09-216-003A-45
; Sequence 45, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Frudakis, Tony N.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462
;  CURRENT APPLICATION NUMBER: US/09/216,003A
;  CURRENT FILING DATE:  1998-12-17
;  NUMBER OF SEQ ID NOS:  310
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-216-003A-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
                ||||||||
Db        532 AAGGTTTC 539


RESULT 23
US-09-667-857-45
; Sequence 45, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C5
;  CURRENT APPLICATION NUMBER: US/09/667,857
;  CURRENT FILING DATE:  2000-09-20
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-667-857-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531


RESULT 24
US-10-198-053-45
; Sequence 45, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-45

  Query Match             11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531
```

```
Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 25
US-09-827-271-45
; Sequence 45, Application US/09827271
; Patent No. 6962980
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
;   LENGTH: 585
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-45

  Query Match           11.3%;  Score 248;  DB 3;  Length 585;
  Best Local Similarity  100.0%;  Pred. No. 6e-113;
  Matches  248;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 351

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 411

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 471

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 531

Qy       2139 AAGGTTTC 2146
             ||||||||
Db        532 AAGGTTTC 539


RESULT 26
US-09-397-787-277
; Sequence 277, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
;  APPLICANT: Benson, Darin R.
;  APPLICANT: Lodes, Michael J.
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
;  TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
;  FILE REFERENCE: 210121.466C2
;  CURRENT APPLICATION NUMBER: US/09/397,787
;  CURRENT FILING DATE: 1999-09-16
;  NUMBER OF SEQ ID NOS: 334
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 277
;   LENGTH: 440
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(440)
;   OTHER INFORMATION: n = A,T,C or G
US-09-397-787-277

  Query Match           9.6%;  Score 211;  DB 3;  Length 440;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-94;
  Matches  211;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 60

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 966
              ||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCC 211


RESULT 27
US-10-237-551-80
; Sequence 80, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
;  APPLICANT: Day, Craig H.
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Parsons, Joseph M.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
;  FILE REFERENCE: 210121.538C3
;  CURRENT APPLICATION NUMBER: US/10/237,551
;  CURRENT FILING DATE:  2002-09-06
;  NUMBER OF SEQ ID NOS: 254
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
;   LENGTH: 228
;   TYPE: DNA
;   ORGANISM: Homo sapiens-ubiquitin UL49-HSV-2
US-10-237-551-80

  Query Match           9.4%;  Score 207;  DB 3;  Length 228;
  Best Local Similarity  100.0%;  Pred. No. 1.3e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 81

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 141

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 201

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGG 752
              ||||||||||||||||||||||||||
Db        202 CACCTGGTACTCCGTCTCAGAGGTGGG 228


RESULT 28
US-09-385-982-169/c
; Sequence 169, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
```

```
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 169
;   LENGTH: 609
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(609)
;   OTHER INFORMATION: n = A,T,C or G
US-09-385-982-169

  Query Match             9.4%;  Score 207;  DB 3;  Length 609;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-92;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 305

Qy       1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 245

Qy       1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 185

Qy       1871 GGTGCTCCGTCTCAGAGGTGGGATGCA 1897
             ||||||||||||||||||||||||||||
Db        184 GGTGCTCCGTCTCAGAGGTGGGATGCA 158


RESULT 29
US-09-404-879A-95/c
; Sequence 95, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462C2
;  CURRENT APPLICATION NUMBER: US/09/404,879A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  393
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 95
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-404-879A-95

  Query Match             8.7%;  Score 191;  DB 3;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 30
US-09-338-933-95/c
; Sequence 95, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer Lynn
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: OVARIAN CANCER
;  FILE REFERENCE: 210121.462C1
;  CURRENT APPLICATION NUMBER: US/09/338,933
;  CURRENT FILING DATE: 1999-06-23
;  NUMBER OF SEQ ID NOS: 312
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-338-933-95

  Query Match             8.7%;  Score 191;  DB 3;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 31
US-09-215-681-95/c
; Sequence 95, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: Frudakis, Tony N.
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
;  TITLE OF INVENTION: OF OVARIAN CANCER
;  FILE REFERENCE: 210121.463
;  CURRENT APPLICATION NUMBER: US/09/215,681A
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 310
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

US-09-215-681-95

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 32
US-09-216-003A-95/c
; Sequence 95, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216,003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-216-003A-95

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 33
US-09-667-857-95/c
; Sequence 95, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-667-857-95

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169
```

RESULT 34
US-10-198-053-95/c
; Sequence 95, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198,053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-053-95

Query Match 8.7%; Score 191; DB 3; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.1e-84;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240
```

```
Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 35
US-09-827-271-95/c
; Sequence 95, Application US/09827271
; Patent No. 6962980
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-95

  Query Match             8.7%;  Score 191;  DB 3;  Length 605;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 300

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 240

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 180

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        179 AAAGAGTCCAC 169


RESULT 36
US-09-404-879A-100
; Sequence 100, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462C2
;  CURRENT APPLICATION NUMBER: US/09/404,879A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  393
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-404-879A-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 37
US-09-338-933-100
; Sequence 100, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer Lynn
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: OVARIAN CANCER
;  FILE REFERENCE: 210121.462C1
;  CURRENT APPLICATION NUMBER: US/09/338,933
;  CURRENT FILING DATE: 1999-06-23
;  NUMBER OF SEQ ID NOS: 312
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-338-933-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 38
US-09-215-681-100
; Sequence 100, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: Frudakis, Tony N.
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
;  TITLE OF INVENTION: OF OVARIAN CANCER
;  FILE REFERENCE: 210121.463
;  CURRENT APPLICATION NUMBER: US/09/215,681A
;  CURRENT FILING DATE: 1998-12-17
```

```
;  NUMBER OF SEQ ID NOS: 310
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-215-681-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 39
US-09-216-003A-100
; Sequence 100, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Frudakis, Tony N.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462
;  CURRENT APPLICATION NUMBER: US/09/216,003A
;  CURRENT FILING DATE:  1998-12-17
;  NUMBER OF SEQ ID NOS:  310
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-216-003A-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 40
US-09-667-857-100
; Sequence 100, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C5
;  CURRENT APPLICATION NUMBER: US/09/667,857
;  CURRENT FILING DATE:  2000-09-20
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-667-857-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
             ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 41
US-10-198-053-100
; Sequence 100, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
              ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 42
US-09-827-271-100
; Sequence 100, Application US/09827271
; Patent No. 6962980
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 697
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-100

  Query Match             8.7%;  Score 191;  DB 3;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-84;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 401

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 461

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 521

Qy       2079 AAAGAGTCCAC 2089
              ||||||||||||
Db        522 AAAGAGTCCAC 532


RESULT 43
US-09-404-879A-99/c
; Sequence 99, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462C2
;  CURRENT APPLICATION NUMBER: US/09/404,879A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  393
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-404-879A-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
              |||
Db         459 GAA 457


RESULT 44
US-09-338-933-99/c
; Sequence 99, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer Lynn
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
;  TITLE OF INVENTION: OVARIAN CANCER
;  FILE REFERENCE: 210121.462C1
;  CURRENT APPLICATION NUMBER: US/09/338,933
;  CURRENT FILING DATE: 1999-06-23
;  NUMBER OF SEQ ID NOS: 312
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-338-933-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
              |||
Db         459 GAA 457


RESULT 45
US-09-215-681-99/c
; Sequence 99, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: Frudakis, Tony N.
;  APPLICANT: King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
;  TITLE OF INVENTION: OF OVARIAN CANCER
;  FILE REFERENCE: 210121.463
;  CURRENT APPLICATION NUMBER: US/09/215,681A
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 310
;  SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-215-681-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 46
US-09-216-003A-99/c
; Sequence 99, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Frudakis, Tony N.
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462
;  CURRENT APPLICATION NUMBER: US/09/216,003A
;  CURRENT FILING DATE:  1998-12-17
;  NUMBER OF SEQ ID NOS:  310
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-216-003A-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 47
US-09-667-857-99/c
; Sequence 99, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C5
;  CURRENT APPLICATION NUMBER: US/09/667,857
;  CURRENT FILING DATE:  2000-09-20
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-667-857-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 48
US-10-198-053-99/c
; Sequence 99, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-198-053-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db        459 GAA 457


RESULT 49
US-09-827-271-99/c
; Sequence 99, Application US/09827271
```

```
; Patent No. 6962980
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 99
;   LENGTH: 636
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-827-271-99

  Query Match             5.6%;  Score 123;  DB 3;  Length 636;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-51;
  Matches  123;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 520

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         519 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 460

Qy       1508 GAA 1510
             |||
Db         459 GAA 457


RESULT 50
US-08-592-126-140/c
; Sequence 140, Application US/08592126
; Patent No. 5821091
;  GENERAL INFORMATION:
;    APPLICANT:  Gregory Dolganov
;    TITLE OF INVENTION:  Transcripts Encoding Immunomodulatory
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  151
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,126
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sholtz, Charles K.
;      REGISTRATION NUMBER:  38,615
;      REFERENCE/DOCKET NUMBER:  4600-0111
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  795 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  E67.seq
US-08-592-126-140

  Query Match             4.7%;  Score 104;  DB 2;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-41;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
             |||||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6


RESULT 51
US-09-168-595-140/c
; Sequence 140, Application US/09168595
; Patent No. 6555666
;  GENERAL INFORMATION:
;    APPLICANT:  Gregory Dolganov
;    TITLE OF INVENTION:  Transcripts Encoding Immunomodulatory
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  151
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/168,595
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,126
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sholtz, Charles K.
;      REGISTRATION NUMBER:  38,615
;      REFERENCE/DOCKET NUMBER:  4600-0111
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  795 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  E67.seq
US-09-168-595-140

  Query Match             4.7%;  Score 104;  DB 3;  Length 795;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-41;
  Matches  104;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        109 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 50

Qy       1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1635
              ||||||||||||||||||||||||||||||||||||||||||||
Db         49 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 6
```

Search completed: March 6, 2006, 23:57:14
Job time : 414 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 6, 2006, 20:42:48 ; Search time 8809 Seconds
(without alignments)
11642.330 Million cell updates/sec

Title: US-10-684-422-1
Perfect score: 2192
Sequence: 1 tagttccgtcgcagccggga.....aataaagttgttgcattccc 2192

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 100

Total number of hits satisfying chosen parameters: 5461

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1890	86.2	2162	4	CR621420	CR621420 full-leng
2	1067	48.7	1201	4	CR611585	CR611585 full-leng
3	918	41.9	1039	1	AL539566	AL539566 AL539566
4	914	41.7	1032	1	AL557436	AL557436 AL557436
5	913	41.7	913	1	AL536209	AL536209 AL536209
6	902	41.1	950	1	AL535635	AL535635 AL535635
7	882	40.2	939	7	CR990253	CR990253 CR990253
8	878	40.1	923	1	AL535723	AL535723 AL535723
9	875	39.9	935	5	BQ955883	BQ955883 AGENCOURT
10	865	39.5	865	1	AL560383	AL560383 AL560383
11	865	39.5	967	4	CR615117	CR615117 full-leng
12	861	39.3	919	5	BU857118	BU857118 AGENCOURT
13	860	39.2	895	7	CR992464	CR992464 CR992464
14	858	39.1	1024	1	AL533035	AL533035 AL533035
15	857	39.1	905	7	CT004068	CT004068 CT004068
16	851	38.8	886	7	CR993134	CR993134 CR993134
17	850	38.8	867	8	CX865290	CX865290 HESC4_2_A
18	845	38.5	845	1	AL535099	AL535099 AL535099
19	843	38.5	843	8	CX873614	CX873614 HESC4_83_
20	842	38.4	881	7	CR997691	CR997691 CR997691
21	841	38.4	937	1	AL560428	AL560428 AL560428
22	840	38.3	847	5	BX419560	BX419560 BX419560
23	836	38.1	915	7	CR981232	CR981232 CR981232
c 24	833	38.0	1026	3	BQ073574	BQ073574 AGENCOURT
25	832	38.0	1104	1	AL535209	AL535209 AL535209
26	830	37.9	913	7	CR998818	CR998818 CR998818
27	829	37.8	829	8	CX871103	CX871103 HESC4_53_
28	827	37.7	839	7	CR996018	CR996018 CR996018
29	825	37.6	831	8	CX866263	CX866263 HESC4_9_E
30	824	37.6	824	7	CR990940	CR990940 CR990940
31	820	37.4	902	6	CD517425	CD517425 AGENCOURT
32	819	37.4	1086	3	BM541766	BM541766 AGENCOURT
33	818	37.3	1009	3	BM450657	BM450657 AGENCOURT
34	814	37.1	885	5	BQ439334	BQ439334 AGENCOURT
35	814	37.1	939	1	AL541568	AL541568 AL541568
36	813	37.1	897	5	BU153047	BU153047 AGENCOURT
37	812	37.0	1009	5	BX458876	BX458876 BX458876
38	811	37.0	831	7	CR990631	CR990631 CR990631
39	810	37.0	819	8	CX865249	CX865249 HESC4_1_F
40	810	37.0	1036	3	BM928821	BM928821 AGENCOURT
41	809	36.9	898	7	CR999015	CR999015 CR999015
42	807	36.8	824	8	CX163519	CX163519 HESC2_5_H
43	807	36.8	895	3	BM551807	BM551807 AGENCOURT
44	803	36.6	889	6	CD244749	CD244749 AGENCOURT
45	803	36.6	919	5	BU187089	BU187089 AGENCOURT
46	801	36.5	801	8	DR762064	DR762064 HESC4_136
47	801	36.5	1022	1	AL555014	AL555014 AL555014
48	801	36.5	1042	1	AL537821	AL537821 AL537821
49	800	36.5	996	1	AL560244	AL560244 AL560244
50	799	36.5	853	7	CR984609	CR984609 CR984609
51	797	36.4	859	7	CR994031	CR994031 CR994031
52	797	36.4	965	1	AL535900	AL535900 AL535900
53	796	36.3	876	5	BU194577	BU194577 AGENCOURT
54	795	36.3	824	7	CR998795	CR998795 CR998795
55	795	36.3	873	3	BQ225402	BQ225402 AGENCOURT
56	795	36.3	879	6	CF552717	CF552717 AGENCOURT
57	794	36.2	794	8	CX866677	CX866677 HESC4_12_
58	794	36.2	945	5	BQ889352	BQ889352 AGENCOURT
59	792	36.1	963	3	BM472268	BM472268 AGENCOURT
60	791	36.1	842	7	CR996412	CR996412 CR996412
61	791	36.1	891	7	CR977355	CR977355 CR977355
62	788	35.9	882	7	CR983573	CR983573 CR983573
63	788	35.9	915	7	CT001833	CT001833 CT001833
64	786	35.9	794	8	CX872472	CX872472 HESC4_69_
65	786	35.9	842	7	CR987122	CR987122 CR987122
66	786	35.9	1037	1	AL535583	AL535583 AL535583
67	785	35.8	836	7	CR987992	CR987992 CR987992
68	785	35.8	872	7	CR995740	CR995740 CR995740
69	784	35.8	903	7	CR987948	CR987948 CR987948
70	784	35.8	905	7	CR991185	CR991185 CR991185
71	784	35.8	973	1	AL535866	AL535866 AL535866
72	784	35.8	1052	3	BM908271	BM908271 AGENCOURT
73	783	35.7	906	5	BU153181	BU153181 AGENCOURT
74	783	35.7	975	3	BQ225018	BQ225018 AGENCOURT
75	782	35.7	825	6	CD244277	CD244277 AGENCOURT
76	782	35.7	859	5	BQ421336	BQ421336 AGENCOURT
77	781	35.6	781	8	CX163961	CX163961 HESC2_11_
78	781	35.6	799	7	CT001360	CT001360 CT001360
79	781	35.6	880	7	CR988748	CR988748 CR988748
80	780	35.6	857	5	BU155752	BU155752 AGENCOURT
81	780	35.6	875	5	BU180938	BU180938 AGENCOURT
82	780	35.6	924	5	BQ880781	BQ880781 AGENCOURT
83	779	35.5	854	5	BU149754	BU149754 AGENCOURT
84	779	35.5	897	1	AL560464	AL560464 AL560464
85	779	35.5	899	5	BU172945	BU172945 AGENCOURT
86	776	35.4	906	7	CR979926	CR979926 CR979926
87	775	35.4	919	5	BQ927751	BQ927751 AGENCOURT
88	774	35.3	920	5	BQ643348	BQ643348 AGENCOURT
89	774	35.3	1070	3	BM463129	BM463129 AGENCOURT
90	773	35.3	773	8	CX871255	CX871255 HESC4_55_
91	773	35.3	1098	5	BU902337	BU902337 AGENCOURT
92	771	35.2	905	5	BQ431211	BQ431211 AGENCOURT
93	771	35.2	926	7	CR978349	CR978349 CR978349
94	771	35.2	1038	3	BQ051425	BQ051425 AGENCOURT
95	770	35.1	770	8	CX873532	CX873532 HESC4_82_

	96	769	35.1	897	5	BU175290	BU175290 AGENCOURT
	97	767	35.0	878	5	BU521200	BU521200 AGENCOURT
	98	767	35.0	917	5	BQ879588	BQ879588 AGENCOURT
	99	767	35.0	977	5	BQ723637	BQ723637 AGENCOURT
	100	767	35.0	1042	3	BM803902	BM803902 AGENCOURT
	101	767	35.0	1080	3	BM464473	BM464473 AGENCOURT
	102	766	34.9	908	5	BQ878048	BQ878048 AGENCOURT
	103	766	34.9	924	5	BQ880336	BQ880336 AGENCOURT
	104	765	34.9	897	7	CR996383	CR996383 CR996383
	105	765	34.9	900	5	BQ894538	BQ894538 AGENCOURT
	106	763	34.8	892	6	CF552130	CF552130 AGENCOURT
	107	763	34.8	911	3	BQ226311	BQ226311 AGENCOURT
	108	763	34.8	1133	3	BM458528	BM458528 AGENCOURT
	109	761	34.7	761	8	DR763328	DR763328 HESC4_151
	110	761	34.7	889	6	CD518820	CD518820 AGENCOURT
	111	761	34.7	908	7	CR997819	CR997819 CR997819
	112	760	34.7	760	8	CX871258	CX871258 HESC4_55_
	113	760	34.7	908	7	CT004188	CT004188 CT004188
	114	760	34.7	917	7	CR999605	CR999605 CR999605
	115	760	34.7	1081	5	BX438063	BX438063 BX438063
	116	759	34.6	862	6	CF552441	CF552441 AGENCOURT
	117	759	34.6	1030	5	BX418049	BX418049 BX418049
	118	759	34.6	1057	3	BM557788	BM557788 AGENCOURT
	119	758	34.6	880	7	CR987157	CR987157 CR987157
	120	758	34.6	931	7	CR976288	CR976288 CR976288
	121	757	34.5	861	5	BU160892	BU160892 AGENCOURT
	122	755	34.4	908	7	CT000802	CT000802 CT000802
	123	755	34.4	922	7	CT001300	CT001300 CT001300
	124	754	34.4	859	6	CD105734	CD105734 AGENCOURT
	125	753	34.4	909	7	CR984559	CR984559 CR984559
	126	753	34.4	927	7	CT001212	CT001212 CT001212
	127	752	34.3	753	8	CX867108	CX867108 HESC4_15_
	128	752	34.3	952	5	BQ888883	BQ888883 AGENCOURT
	129	751	34.3	845	6	CD244644	CD244644 AGENCOURT
	130	751	34.3	855	5	BQ881766	BQ881766 AGENCOURT
	131	750	34.2	750	8	DR762082	DR762082 HESC4_136
	132	750	34.2	1114	3	BM455235	BM455235 AGENCOURT
	133	749	34.2	910	7	CR985228	CR985228 CR985228
	134	748	34.1	768	6	CF131969	CF131969 UI-HF-FQ0
	135	748	34.1	817	6	CF131877	CF131877 UI-HF-FQ0
	136	748	34.1	899	3	BQ233859	BQ233859 AGENCOURT
	137	748	34.1	1004	3	BM550155	BM550155 AGENCOURT
	138	748	34.1	1074	5	BX358407	BX358407 BX358407
	139	747	34.1	995	3	BQ050835	BQ050835 AGENCOURT
c	140	746	34.0	774	1	AW409943	AW409943 fh03c07.y
	141	746	34.0	878	6	CF552282	CF552282 AGENCOURT
	142	746	34.0	884	3	BQ215503	BQ215503 AGENCOURT
	143	745	34.0	938	7	CT001459	CT001459 CT001459
	144	744	33.9	744	8	CX786633	CX786633 HESC3_62_
	145	744	33.9	753	6	CF131596	CF131596 UI-HF-FQ0
	146	744	33.9	764	6	CF132751	CF132751 UI-HF-FQ0
	147	743	33.9	766	8	CX787889	CX787889 HESC3_91_
	148	743	33.9	768	6	CF132418	CF132418 UI-HF-FQ0
	149	743	33.9	782	8	DR759783	DR759783 HESC4_107
	150	743	33.9	813	6	CF132037	CF132037 UI-HF-FQ0
	151	743	33.9	912	7	CT001169	CT001169 CT001169
	152	743	33.9	1135	3	BM462979	BM462979 AGENCOURT
c	153	742	33.9	760	3	BM999891	BM999891 UI-H-DP0-
	154	742	33.9	780	6	CF131371	CF131371 UI-HF-FQ0
	155	742	33.9	808	6	CF132239	CF132239 UI-HF-FQ0
	156	742	33.9	857	5	BQ879520	BQ879520 AGENCOURT
	157	742	33.9	882	6	CF552062	CF552062 AGENCOURT
	158	742	33.9	882	6	CF552518	CF552518 AGENCOURT
	159	741	33.8	920	5	BQ722449	BQ722449 AGENCOURT
	160	740	33.8	958	5	BQ934306	BQ934306 AGENCOURT
	161	738	33.7	893	7	CR983401	CR983401 CR983401
	162	738	33.7	897	3	BQ226817	BQ226817 AGENCOURT
	163	737	33.6	737	1	AL536994	AL536994 AL536994
	164	737	33.6	859	6	CD519152	CD519152 AGENCOURT
	165	737	33.6	945	5	BU182397	BU182397 AGENCOURT
	166	734	33.5	734	8	DR761274	DR761274 HESC4_126
	167	734	33.5	888	6	CD517713	CD517713 AGENCOURT
	168	734	33.5	915	7	CR978764	CR978764 CR978764
	169	734	33.5	982	5	BU146473	BU146473 AGENCOURT
	170	733	33.4	733	7	CR991559	CR991559 CR991559
	171	733	33.4	917	5	BQ948667	BQ948667 AGENCOURT
	172	733	33.4	1074	1	AL541647	AL541647 AL541647
	173	732	33.4	902	5	BU165709	BU165709 AGENCOURT
	174	732	33.4	925	7	CR977715	CR977715 CR977715
	175	731	33.3	731	8	CX873035	CX873035 HESC4_76_
	176	731	33.3	738	7	CN415690	CN415690 170006002
	177	731	33.3	854	6	CD517104	CD517104 AGENCOURT
	178	731	33.3	877	3	BQ224093	BQ224093 AGENCOURT
	179	731	33.3	925	6	CF552082	CF552082 AGENCOURT
	180	731	33.3	936	6	CD557608	CD557608 AGENCOURT
	181	730	33.3	739	7	CT000208	CT000208 CT000208
	182	730	33.3	834	6	CD245746	CD245746 AGENCOURT
	183	730	33.3	890	6	CD517028	CD517028 AGENCOURT
	184	730	33.3	908	7	CT003566	CT003566 CT003566
	185	730	33.3	970	3	BM563395	BM563395 AGENCOURT
	186	729	33.3	729	8	CX164520	CX164520 HESC2_17_
	187	729	33.3	729	8	CX873731	CX873731 HESC4_84_
	188	729	33.3	729	8	DN601704	DN601704 HESC3_71_
	189	729	33.3	729	8	DR759749	DR759749 HESC4_107
	190	729	33.3	729	8	DR763210	DR763210 HESC4_149
c	191	729	33.3	780	1	AL603620	AL603620 DKFZp434K
	192	729	33.3	879	3	BQ212090	BQ212090 AGENCOURT
	193	729	33.3	893	3	BQ227353	BQ227353 AGENCOURT
	194	729	33.3	1053	3	BM909385	BM909385 AGENCOURT
	195	728	33.2	865	7	CR989749	CR989749 CR989749
	196	727	33.2	727	8	CX868946	CX868946 HESC4_31_
	197	727	33.2	857	5	BU178242	BU178242 AGENCOURT
	198	727	33.2	861	7	CR984951	CR984951 CR984951
	199	727	33.2	864	6	CD517161	CD517161 AGENCOURT
	200	726	33.1	741	7	CR999018	CR999018 CR999018
	201	726	33.1	773	7	CR983470	CR983470 CR983470
	202	726	33.1	1088	3	BM557281	BM557281 AGENCOURT
	203	725	33.1	835	6	CD514337	CD514337 AGENCOURT
	204	725	33.1	904	5	BU172676	BU172676 AGENCOURT
	205	724	33.0	724	8	CX164590	CX164590 HESC2_18_
	206	724	33.0	724	8	CX866221	CX866221 HESC4_9_B
	207	724	33.0	724	8	CX873220	CX873220 HESC4_78_
	208	724	33.0	742	7	CN415660	CN415660 170004241
	209	724	33.0	849	7	CR993103	CR993103 CR993103
	210	724	33.0	875	3	BQ220564	BQ220564 AGENCOURT
	211	724	33.0	883	5	BU153538	BU153538 AGENCOURT
	212	724	33.0	900	5	BQ434497	BQ434497 AGENCOURT
	213	724	33.0	905	5	BU195402	BU195402 AGENCOURT
	214	724	33.0	927	5	BQ670207	BQ670207 AGENCOURT
c	215	723	33.0	889	5	BX414032	BX414032 BX414032
	216	722	32.9	860	5	BU151351	BU151351 AGENCOURT
	217	721	32.9	789	7	CT000659	CT000659 CT000659
	218	720	32.8	736	8	DR422128	DR422128 nav08d04.
	219	720	32.8	813	5	BQ961879	BQ961879 AGENCOURT
	220	720	32.8	910	5	BQ935362	BQ935362 AGENCOURT
	221	719	32.8	719	8	CX165664	CX165664 HESC2_31_
	222	719	32.8	719	8	DR155702	DR155702 HESC2_57_
	223	718	32.8	769	8	DR759813	DR759813 HESC4_108
	224	718	32.8	865	5	BU164691	BU164691 AGENCOURT
	225	718	32.8	884	5	BU171951	BU171951 AGENCOURT
	226	718	32.8	894	6	CD251905	CD251905 AGENCOURT
	227	718	32.8	1008	5	BX442070	BX442070 BX442070
	228	717	32.7	717	7	CN415659	CN415659 170005829
	229	717	32.7	725	7	CN415629	CN415629 170005999
	230	717	32.7	897	6	CD556735	CD556735 AGENCOURT
	231	717	32.7	970	5	BU175009	BU175009 AGENCOURT
	232	716	32.7	716	6	CF131936	CF131936 UI-HF-FQ0
	233	716	32.7	716	7	CT004302	CT004302 CT004302
	234	716	32.7	716	8	CX872917	CX872917 HESC4_74_
	235	716	32.7	816	5	BU193795	BU193795 AGENCOURT
	236	716	32.7	888	5	BU150027	BU150027 AGENCOURT
	237	716	32.7	920	7	CT003348	CT003348 CT003348
	238	716	32.7	930	7	CR988580	CR988580 CR988580
	239	716	32.7	936	7	CR977118	CR977118 CR977118
	240	716	32.7	949	5	BQ889973	BQ889973 AGENCOURT
	241	716	32.7	1006	3	BQ049067	BQ049067 AGENCOURT

	242	716	32.7	1093	3	BM477326	BM477326 AGENCOURT
	243	715	32.6	1070	3	BM476508	BM476508 AGENCOURT
	244	715	32.6	1102	1	AL516880	AL516880 AL516880
	245	714	32.6	714	8	CX873027	CX873027 HESC4_75_
	246	713	32.5	906	5	BQ934245	BQ934245 AGENCOURT
	247	712	32.5	842	1	AU130085	AU130085 AU130085
	248	712	32.5	875	5	BQ421393	BQ421393 AGENCOURT
	249	712	32.5	875	6	CD515652	CD515652 AGENCOURT
	250	711	32.4	711	8	DR156196	DR156196 HESC2_63_
	251	711	32.4	711	8	DR760405	DR760405 HESC4_115
	252	711	32.4	762	8	CX782198	CX782198 HESC3_12_
	253	710	32.4	942	5	BQ879740	BQ879740 AGENCOURT
	254	709	32.3	851	5	BQ881494	BQ881494 AGENCOURT
c	255	708	32.3	726	3	BM975559	BM975559 UI-CF-EN1
	256	708	32.3	769	6	CF132206	CF132206 UI-HF-FQ0
c	257	708	32.3	915	5	BU557860	BU557860 AGENCOURT
	258	707	32.3	783	6	CF131782	CF131782 UI-HF-FQ0
	259	707	32.3	885	5	BQ689319	BQ689319 AGENCOURT
	260	707	32.3	896	7	CR982652	CR982652 CR982652
	261	707	32.3	911	7	CR985699	CR985699 CR985699
	262	707	32.3	913	7	CT002282	CT002282 CT002282
	263	707	32.3	916	5	BQ876710	BQ876710 AGENCOURT
	264	706	32.2	842	5	BU178105	BU178105 AGENCOURT
	265	706	32.2	842	7	CR980550	CR980550 CR980550
	266	706	32.2	892	6	CD556696	CD556696 AGENCOURT
	267	705	32.2	872	6	CD106257	CD106257 AGENCOURT
	268	705	32.2	963	5	BQ898339	BQ898339 AGENCOURT
	269	704	32.1	704	7	CN415631	CN415631 170006001
	270	704	32.1	704	8	CX873406	CX873406 HESC4_80_
	271	704	32.1	718	8	CX868172	CX868172 HESC4_26_
	272	704	32.1	729	6	CF132345	CF132345 UI-HF-FQ0
	273	704	32.1	773	6	CF131892	CF131892 UI-HF-FQ0
	274	704	32.1	779	6	CF132167	CF132167 UI-HF-FQ0
	275	704	32.1	859	5	BU175603	BU175603 AGENCOURT
	276	703	32.1	703	8	CX873206	CX873206 HESC4_78_
	277	703	32.1	747	6	CF131818	CF131818 UI-HF-FQ0
	278	703	32.1	1091	3	BM455749	BM455749 AGENCOURT
	279	702	32.0	855	5	BU859982	BU859982 AGENCOURT
	280	702	32.0	896	3	BQ222115	BQ222115 AGENCOURT
	281	702	32.0	913	5	BU173376	BU173376 AGENCOURT
	282	702	32.0	929	5	BQ717647	BQ717647 AGENCOURT
	283	702	32.0	949	5	BQ719762	BQ719762 AGENCOURT
	284	700	31.9	847	5	BU159353	BU159353 AGENCOURT
	285	700	31.9	900	5	BQ877781	BQ877781 AGENCOURT
	286	700	31.9	956	5	BU146224	BU146224 AGENCOURT
	287	699	31.9	866	5	BX460542	BX460542 BX460542
	288	699	31.9	869	5	BQ897615	BQ897615 AGENCOURT
	289	699	31.9	908	5	BQ927395	BQ927395 AGENCOURT
	290	698	31.8	755	7	CN415661	CN415661 170006000
	291	698	31.8	893	1	AL536406	AL536406 AL536406
	292	697	31.8	785	3	BQ216161	BQ216161 AGENCOURT
	293	697	31.8	826	5	BU195276	BU195276 AGENCOURT
	294	696	31.8	993	3	BQ050542	BQ050542 AGENCOURT
	295	695	31.7	893	7	CR984160	CR984160 CR984160
c	296	693	31.6	711	3	BQ181370	BQ181370 UI-H-EU0-
	297	693	31.6	778	6	CF132391	CF132391 UI-HF-FQ0
	298	693	31.6	831	7	CR992640	CR992640 CR992640
	299	693	31.6	841	5	BU162203	BU162203 AGENCOURT
	300	693	31.6	873	7	CR978503	CR978503 CR978503
	301	692	31.6	692	8	CX165746	CX165746 HESC2_32_
	302	692	31.6	792	6	CF456877	CF456877 AGENCOURT
	303	692	31.6	807	1	AU142670	AU142670 AU142670
	304	692	31.6	814	6	CF132042	CF132042 UI-HF-FQ0
c	305	691	31.5	760	3	BQ181836	BQ181836 UI-H-EU0-
	306	691	31.5	1087	1	AL514142	AL514142 AL514142
	307	690	31.5	741	8	CX867120	CX867120 HESC4_15_
	308	690	31.5	825	1	AU132831	AU132831 AU132831
	309	690	31.5	884	5	BU165811	BU165811 AGENCOURT
	310	690	31.5	926	3	BQ228219	BQ228219 AGENCOURT
	311	689	31.4	1085	5	BQ436578	BQ436578 AGENCOURT
	312	688	31.4	759	6	CD103348	CD103348 AGENCOURT
	313	688	31.4	793	6	CF128312	CF128312 UI-HF-ET0
	314	688	31.4	856	7	CR986382	CR986382 CR986382
	315	687	31.3	742	6	CF132417	CF132417 UI-HF-FQ0
	316	687	31.3	756	7	CR559535	CR559535 DKFZp468L
	317	686	31.3	848	5	BQ925272	BQ925272 AGENCOURT
	318	686	31.3	933	5	BU899304	BU899304 AGENCOURT
	319	685	31.2	867	7	CR994423	CR994423 CR994423
	320	685	31.2	889	5	BU156142	BU156142 AGENCOURT
	321	684	31.2	870	6	CD244139	CD244139 AGENCOURT
	322	684	31.2	943	6	CD516062	CD516062 AGENCOURT
	323	684	31.2	1084	3	BM467727	BM467727 AGENCOURT
	324	683	31.2	883	5	BQ668779	BQ668779 AGENCOURT
	325	683	31.2	895	7	CR979111	CR979111 CR979111
c	326	682	31.1	700	6	CA414684	CA414684 UI-H-EZ0-
	327	682	31.1	742	8	CX757996	CX757996 AGENCOURT
	328	682	31.1	868	5	BU196155	BU196155 AGENCOURT
	329	682	31.1	899	5	BQ896764	BQ896764 AGENCOURT
	330	681	31.1	782	1	AU126027	AU126027 AU126027
	331	681	31.1	866	5	BU150096	BU150096 AGENCOURT
	332	681	31.1	889	7	CR998279	CR998279 CR998279
	333	681	31.1	925	7	CR977834	CR977834 CR977834
	334	681	31.1	1021	3	BM907782	BM907782 AGENCOURT
	335	680	31.0	745	7	CR993630	CR993630 CR993630
	336	680	31.0	827	3	BQ213115	BQ213115 AGENCOURT
	337	680	31.0	839	5	BU172726	BU172726 AGENCOURT
	338	680	31.0	913	5	BQ717223	BQ717223 AGENCOURT
	339	680	31.0	1001	3	BQ051270	BQ051270 AGENCOURT
	340	680	31.0	1057	1	AL559920	AL559920 AL559920
	341	679	31.0	865	6	CD245048	CD245048 AGENCOURT
	342	679	31.0	877	6	CD519290	CD519290 AGENCOURT
	343	679	31.0	886	7	CR982284	CR982284 CR982284
	344	679	31.0	905	7	CR985573	CR985573 CR985573
	345	678	30.9	773	7	CN415655	CN415655 170005999
	346	678	30.9	864	5	BQ439676	BQ439676 AGENCOURT
	347	678	30.9	1099	3	BM552888	BM552888 AGENCOURT
	348	677	30.9	677	8	CX868155	CX868155 HESC4_26_
	349	677	30.9	719	7	CR990546	CR990546 CR990546
	350	677	30.9	729	6	CF131445	CF131445 UI-HF-FQ0
	351	677	30.9	735	8	CX165200	CX165200 HESC2_25_
	352	677	30.9	799	1	AU124338	AU124338 AU124338
	353	677	30.9	945	5	BQ956984	BQ956984 AGENCOURT
	354	677	30.9	1092	1	AL538959	AL538959 AL538959
	355	676	30.8	732	1	AU143561	AU143561 AU143561
	356	676	30.8	938	5	BQ890535	BQ890535 AGENCOURT
	357	674	30.7	676	7	CR996041	CR996041 CR996041
	358	674	30.7	706	6	CB957633	CB957633 AGENCOURT
	359	674	30.7	716	8	CX870578	CX870578 HESC4_47_
	360	674	30.7	784	1	AU141404	AU141404 AU141404
	361	674	30.7	881	5	BQ894565	BQ894565 AGENCOURT
c	362	673	30.7	742	3	BQ000285	BQ000285 UI-H-DP0-
	363	673	30.7	845	7	CR995922	CR995922 CR995922
	364	673	30.7	925	5	BQ675781	BQ675781 AGENCOURT
	365	673	30.7	955	5	BQ686784	BQ686784 AGENCOURT
	366	672	30.7	672	7	CN415711	CN415711 170005830
	367	672	30.7	911	7	CR996695	CR996695 CR996695
	368	671	30.6	832	3	BQ224141	BQ224141 AGENCOURT
	369	671	30.6	935	5	BQ894244	BQ894244 AGENCOURT
	370	670	30.6	756	7	CN415619	CN415619 170005999
	371	670	30.6	862	5	BQ421791	BQ421791 AGENCOURT
	372	669	30.5	673	6	CF131403	CF131403 UI-HF-FQ0
	373	669	30.5	768	5	BQ433457	BQ433457 AGENCOURT
	374	669	30.5	856	5	BQ721025	BQ721025 AGENCOURT
	375	669	30.5	919	5	BQ892927	BQ892927 AGENCOURT
	376	669	30.5	966	5	BU508547	BU508547 AGENCOURT
	377	669	30.5	1065	3	BM551765	BM551765 AGENCOURT
	378	668	30.5	842	5	BU183761	BU183761 AGENCOURT
	379	668	30.5	968	5	BU509473	BU509473 AGENCOURT
	380	668	30.5	1012	3	BM472131	BM472131 AGENCOURT
	381	668	30.5	1117	5	BU902021	BU902021 AGENCOURT
	382	667	30.4	858	7	CT001569	CT001569 CT001569
	383	667	30.4	864	5	BQ436844	BQ436844 AGENCOURT
	384	667	30.4	906	5	BU180891	BU180891 AGENCOURT
	385	667	30.4	910	3	BQ232136	BQ232136 AGENCOURT
	386	666	30.4	848	3	BQ216706	BQ216706 AGENCOURT
	387	666	30.4	902	7	CR982098	CR982098 CR982098

	388	665	30.3	724	7	CN415709	CN415709 170006001
	389	665	30.3	899	7	CR979224	CR979224 CR979224
	390	665	30.3	1086	5	BX417072	BX417072 BX417072
	391	664	30.3	757	6	CF132101	CF132101 UI-HF-FQ0
	392	664	30.3	802	1	AU121495	AU121495 AU121495
	393	664	30.3	819	7	CR980496	CR980496 CR980496
	394	664	30.3	922	5	BQ425647	BQ425647 AGENCOURT
	395	663	30.2	671	7	CN415696	CN415696 170005999
	396	663	30.2	691	7	CN415625	CN415625 170004241
	397	663	30.2	706	6	CF131533	CF131533 UI-HF-FQ0
	398	663	30.2	963	5	BQ716598	BQ716598 AGENCOURT
	399	663	30.2	1033	3	BM469668	BM469668 AGENCOURT
	400	662	30.2	734	7	CN415628	CN415628 170006000
	401	662	30.2	833	5	BQ893358	BQ893358 AGENCOURT
	402	662	30.2	864	5	BQ721570	BQ721570 AGENCOURT
	403	662	30.2	884	3	BQ221580	BQ221580 AGENCOURT
	404	662	30.2	899	1	AL546275	AL546275 AL546275
	405	662	30.2	906	5	BQ723374	BQ723374 AGENCOURT
	406	662	30.2	906	5	BQ960775	BQ960775 AGENCOURT
	407	660	30.1	741	1	AU124586	AU124586 AU124586
	408	660	30.1	791	1	AU119145	AU119145 AU119145
	409	658	30.0	746	6	CD511509	CD511509 AGENCOURT
	410	657	30.0	657	8	CX869712	CX869712 HESC4_36_
c	411	657	30.0	725	6	CA443858	CA443858 UI-H-DP0-
c	412	657	30.0	726	3	BQ183386	BQ183386 UI-H-EU0-
c	413	657	30.0	751	3	BQ000037	BQ000037 UI-H-DP0-
	414	657	30.0	800	3	BQ225717	BQ225717 AGENCOURT
	415	657	30.0	913	5	BQ918091	BQ918091 AGENCOURT
	416	657	30.0	930	6	CF780728	CF780728 AGENCOURT
	417	657	30.0	983	5	BX418982	BX418982 BX418982
	418	656	29.9	669	7	CR998106	CR998106 CR998106
	419	656	29.9	868	5	BQ423460	BQ423460 AGENCOURT
	420	656	29.9	942	5	BQ440023	BQ440023 AGENCOURT
	421	656	29.9	1050	3	BM469063	BM469063 AGENCOURT
	422	655	29.9	692	6	CF132195	CF132195 UI-HF-FQ0
	423	655	29.9	712	7	CR630423	CR630423 DKFZp469L
	424	655	29.9	950	5	BQ669560	BQ669560 AGENCOURT
	425	655	29.9	1007	3	BM561265	BM561265 AGENCOURT
	426	653	29.8	664	7	CK003157	CK003157 AGENCOURT
c	427	653	29.8	685	6	CA445721	CA445721 UI-H-ED0-
	428	653	29.8	704	2	BF727156	BF727156 by17e11.y
	429	653	29.8	751	6	CF131948	CF131948 UI-HF-FQ0
	430	653	29.8	812	6	CD643439	CD643439 AGENCOURT
c	431	653	29.8	867	5	BX434493	BX434493 BX434493
c	432	653	29.8	867	5	BX434832	BX434832 BX434832
	433	653	29.8	876	7	CR979139	CR979139 CR979139
	434	653	29.8	884	6	CD300421	CD300421 AGENCOURT
c	435	653	29.8	888	5	BX455126	BX455126 BX455126
	436	652	29.7	728	6	CF131647	CF131647 UI-HF-FQ0
	437	652	29.7	885	5	BQ880488	BQ880488 AGENCOURT
	438	651	29.7	873	5	BQ676274	BQ676274 AGENCOURT
	439	650	29.7	742	5	BQ953412	BQ953412 AGENCOURT
	440	650	29.7	865	5	BU150123	BU150123 AGENCOURT
	441	649	29.6	735	1	AU125985	AU125985 AU125985
	442	649	29.6	759	6	CB986737	CB986737 AGENCOURT
	443	649	29.6	769	1	AU140032	AU140032 AU140032
	444	649	29.6	889	6	CD359147	CD359147 AGENCOURT
	445	649	29.6	895	5	BQ955771	BQ955771 AGENCOURT
c	446	649	29.6	900	5	BX455930	BX455930 BX455930
	447	649	29.6	906	3	BQ221089	BQ221089 AGENCOURT
	448	648	29.6	648	8	CX868187	CX868187 HESC4_26_
	449	648	29.6	882	5	BQ948620	BQ948620 AGENCOURT
	450	648	29.6	1023	3	BM476008	BM476008 AGENCOURT
	451	647	29.5	716	1	AU143379	AU143379 AU143379
	452	647	29.5	738	1	AU143730	AU143730 AU143730
	453	647	29.5	875	6	CD109038	CD109038 AGENCOURT
	454	647	29.5	927	7	CR987568	CR987568 CR987568
	455	647	29.5	976	5	BQ720635	BQ720635 AGENCOURT
	456	646	29.5	739	6	CD101774	CD101774 AGENCOURT
	457	645	29.4	678	7	CR999200	CR999200 CR999200
	458	645	29.4	873	5	BQ720220	BQ720220 AGENCOURT
c	459	645	29.4	883	5	BQ673316	BQ673316 AGENCOURT
	460	644	29.4	655	2	BG682234	BG682234 602629542
	461	644	29.4	695	7	CR975351	CR975351 CR975351
	462	644	29.4	697	7	CN415721	CN415721 170004248
c	463	644	29.4	728	3	BQ008664	BQ008664 UI-H-EI0-
	464	644	29.4	728	7	CN415630	CN415630 170004249
c	465	644	29.4	929	5	BX455770	BX455770 BX455770
	466	643	29.3	661	5	BU194759	BU194759 AGENCOURT
	467	643	29.3	846	5	BU153946	BU153946 AGENCOURT
	468	643	29.3	893	5	BU189314	BU189314 AGENCOURT
	469	643	29.3	903	5	BQ721932	BQ721932 AGENCOURT
	470	643	29.3	908	7	CR995472	CR995472 CR995472
	471	642	29.3	657	6	CF131349	CF131349 UI-HF-FQ0
	472	642	29.3	744	7	CT005219	CT005219 CT005219
	473	642	29.3	745	6	CF131237	CF131237 UI-HF-FQ0
	474	642	29.3	747	6	CF132733	CF132733 UI-HF-FQ0
	475	642	29.3	752	6	CF131521	CF131521 UI-HF-FQ0
	476	642	29.3	754	6	CF132742	CF132742 UI-HF-FQ0
c	477	642	29.3	762	6	CB850720	CB850720 UI-CF-EN1
	478	642	29.3	813	6	CF132251	CF132251 UI-HF-FQ0
	479	642	29.3	814	6	CF132530	CF132530 UI-HF-FQ0
	480	642	29.3	829	5	BU166951	BU166951 AGENCOURT
	481	642	29.3	875	6	CD514210	CD514210 AGENCOURT
	482	642	29.3	901	3	BM449797	BM449797 AGENCOURT
	483	642	29.3	916	6	CD558893	CD558893 AGENCOURT
c	484	641	29.2	659	6	CA412118	CA412118 UI-H-EZ0-
	485	641	29.2	705	7	CN415626	CN415626 170006008
	486	641	29.2	760	6	CF132504	CF132504 UI-HF-FQ0
	487	641	29.2	875	5	BQ432005	BQ432005 AGENCOURT
	488	641	29.2	892	5	BQ722420	BQ722420 AGENCOURT
c	489	641	29.2	1020	5	BQ952050	BQ952050 AGENCOURT
	490	640	29.2	680	8	DR156316	DR156316 HESC2_64_
	491	640	29.2	711	6	CF132680	CF132680 UI-HF-FQ0
	492	640	29.2	793	3	BM466997	BM466997 AGENCOURT
	493	640	29.2	796	8	CX760460	CX760460 AGENCOURT
	494	640	29.2	811	5	BQ430504	BQ430504 AGENCOURT
	495	640	29.2	827	5	BX367020	BX367020 BX367020
	496	640	29.2	939	5	BQ878852	BQ878852 AGENCOURT
	497	639	29.2	663	7	CN415672	CN415672 170004245
	498	639	29.2	1099	3	BM476450	BM476450 AGENCOURT
c	499	638	29.1	638	5	BU786863	BU786863 im94d07.x
	500	638	29.1	883	6	CD300649	CD300649 AGENCOURT

ALIGNMENTS

RESULT 1
CR621420
LOCUS CR621420 2162 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DE004YK08 of Placenta of Homo sapiens (human).
ACCESSION CR621420
VERSION CR621420.1 GI:50502227
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2162)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2162)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES Location/Qualifiers
source 1. .2162
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE004YK08"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 86.2%; Score 1890; DB 4; Length 2162;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2140; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 77

Qy      92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 137

Qy     152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     138 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 197

Qy     212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     198 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 257

Qy     272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     258 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 317

Qy     332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     318 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 377

Qy     392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     378 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 437

Qy     452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     438 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 497

Qy     512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     498 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 557

Qy     572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     558 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 617

Qy     632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     618 TCCTTCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 677

Qy     692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     678 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 737

Qy     752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     738 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 797

Qy     812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 871
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     798 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 857

Qy     872 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 931
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     858 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 917

Qy     932 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 991
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     918 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 977

Qy     992 CGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGA 1051
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     978 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1037

Qy    1052 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1111
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1038 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1097

Qy    1112 CTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGA 1171
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1098 CTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGA 1157

Qy    1172 GTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1231
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1158 GTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1217

Qy    1232 GACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1291
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1218 GACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1277

Qy    1292 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAA 1351
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1278 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1337

Qy    1352 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1338 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1397

Qy    1412 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1471
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1398 CCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1457

Qy    1472 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1531
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1458 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1517

Qy    1532 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1591
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1518 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGA 1577

Qy    1592 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1651
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1578 TGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCG 1637

Qy    1652 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1711
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1638 TCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1697

Qy    1712 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1771
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1698 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1757

Qy    1772 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1831
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1758 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1817

Qy    1832 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 1891
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1818 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 1877

Qy    1892 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 1951
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1878 GATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 1937

Qy    1952 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCA 2011
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1938 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCA 1997

Qy       2012 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 2071
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1998 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 2057

Qy       2072 CATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTC 2131
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2058 CATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGGTGTCTAAGTTTC 2117

Qy       2132 CCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATA 2176
                ||||||||||||||||||||||||||||||||||||||||||||
Db       2118 CCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATA 2162


RESULT 2
CR611585
LOCUS       CR611585               1201 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CL0BB015ZB04 of Neuroblastoma of Homo
            sapiens (human).
ACCESSION   CR611585
VERSION     CR611585.1  GI:50492392
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1201)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CL0BB015ZB04"
                     /tissue_type="Neuroblastoma"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             48.7%;  Score 1067;  DB 4;  Length 1201;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1167;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 92

Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 152

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 212

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 272

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 332

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
                || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 392

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 452

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 512

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 572

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 632

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 692

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 752

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 812

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 872

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1922
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        873 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGT 932

Qy       1923 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 1982
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 992

Qy       1983 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 2042
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        993 CAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 1052

Qy       2043 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 2102
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1053 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTC 1112

Qy       2103 CTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 2162
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1113 CTGCGCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCA 1172

Qy       2163 CTTTCCTTTCAATAAAGTTGTTGCATTCC 2191
                |||||||||||||||||||||||||||||
Db       1173 CTTTCCTTTCAATAAAGTTGTTGCATTCC 1201


RESULT 3
AL539566
LOCUS       AL539566               1039 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL539566 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
```

```
            CS0DF036YO05 5-PRIME, mRNA sequence.
ACCESSION   AL539566
VERSION     AL539566.3  GI:45715224
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1039)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31264133.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF036AH03QP1&c=6328.r.
FEATURES                Location/Qualifiers
     source             1. .1039
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CS0DF036YO05"
                        /tissue_type="FETAL BRAIN"
                        /dev_stage="fetal"
                        /clone_lib="Homo sapiens FETAL BRAIN"
                        /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                        was primed with a NotI-oligo(dT) primer. Five prime end
                        enriched, double-strand cDNA was digested with Not I and
                        cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                        vector. Library was not normalized."
ORIGIN
```

Query Match 41.9%; Score 918; DB 1; Length 1039;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 968; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 67

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 127

Qy        140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 187

Qy        200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 247

Qy        260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 307

Qy        320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 367

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 427

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 487

Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 547

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 607

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 667

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 727

Qy        740 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 787

Qy        800 GGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCAT 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 GGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCAT 847

Qy        860 TCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCT 919
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        848 TCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGRAAGCAGCTGGAAGATGGGCGCACCCT 907

Qy        920 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        908 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 967

Qy        980 GATGCAGAT 988
              |||||||||
Db        968 GATGCAGAT 976
```

```
RESULT 4
AL557436
LOCUS       AL557436                1032 bp    mRNA    linear   EST 02-APR-2004
DEFINITION  AL557436 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
            clone CS0DH003YI04 5-PRIME, mRNA sequence.
ACCESSION   AL557436
VERSION     AL557436.3  GI:46182844
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1032)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31279236.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DH003BE02QP1&c=6328.r.
FEATURES                Location/Qualifiers
```

```
     source          1. .1032
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DH003YI04"
                     /tissue_type="T CELLS (JURKAT CELL LINE)"
                     /cell_line="JURKAT CELL LINE"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             41.7%;  Score 914;  DB 1;  Length 1032;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  914;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 900

Qy        915 ACCCTGTCTGACTA 928
              ||||||||||||||
Db        901 ACCCTGTCTGACTA 914


RESULT 5
AL536209
LOCUS       AL536209                913 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL536209 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF018YI16 5-PRIME, mRNA sequence.
ACCESSION   AL536209
VERSION     AL536209.3  GI:45712045
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31261079.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF018BE08QP1&c=6328.r.
FEATURES              Location/Qualifiers
     source           1. .913
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DF018YI16"
                      /tissue_type="FETAL BRAIN"
                      /dev_stage="fetal"
                      /clone_lib="Homo sapiens FETAL BRAIN"
                      /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                      was primed with a NotI-oligo(dT) primer. Five prime end
                      enriched, double-strand cDNA was digested with Not I and
                      cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                      vector. Library was not normalized."
ORIGIN

  Query Match             41.7%;  Score 913;  DB 1;  Length 913;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  913;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180
```

```
Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CCCTGTCTGACTA 928
              |||||||||||||
Db        901 CCCTGTCTGACTA 913
```

```
RESULT 6
AL535635
LOCUS       AL535635                950 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535635 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF016YG24 5-PRIME, mRNA sequence.
ACCESSION   AL535635
VERSION     AL535635.3  GI:45711502
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 950)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31260684.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF016BD12QP1&c=6328.r.
FEATURES              Location/Qualifiers
     source           1. .950
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DF016YG24"
                      /tissue_type="FETAL BRAIN"
                      /dev_stage="fetal"
                      /clone_lib="Homo sapiens FETAL BRAIN"
                      /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                      was primed with a NotI-oligo(dT) primer. Five prime end
                      enriched, double-strand cDNA was digested with Not I and
                      cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                      vector. Library was not normalized."
ORIGIN
```

```
  Query Match             41.1%;  Score 902;  DB 1;  Length 950;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  902;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660
```

```
Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 900

Qy        916 CC 917
              ||
Db        901 CC 902


RESULT 7
CR990253
LOCUS       CR990253                939 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR990253 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016A1712 5',
            mRNA sequence.
ACCESSION   CR990253
VERSION     CR990253.1  GI:68284138
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016A1712.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016A1712
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016A1712"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             40.2%;  Score 882;  DB 7;  Length 939;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  882;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 60

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 120

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 180

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 240

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 300

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 360

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 420

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 480

Qy        504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 540

Qy        564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 600

Qy        624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 660

Qy        684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 720

Qy        744 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 780

Qy        804 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 840

Qy        864 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 882


RESULT 8
AL535723
LOCUS       AL535723                923 bp    mRNA    linear   EST 24-MAR-2004
```

```
DEFINITION  AL535723 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF016YN11 5-PRIME, mRNA sequence.
ACCESSION   AL535723
VERSION     AL535723.3  GI:45711581
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 923)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31260725.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF016CG06QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .923
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF016YN11"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             40.1%;  Score 878;  DB 1;  Length 923;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  878;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 60

Qy         78 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 120

Qy        138 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 180

Qy        198 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 240

Qy        258 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 300

Qy        318 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 360

Qy        378 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 420

Qy        438 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 480

Qy        498 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 540

Qy        558 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 600

Qy        618 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 660

Qy        678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 720

Qy        738 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 780

Qy        798 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGC 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGC 840

Qy        858 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 895
              |||||||||||||||||||||||||||||||||||||||
Db        841 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 878


RESULT 9
BQ955883
LOCUS       BQ955883                935 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8818082 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6421959
            5', mRNA sequence.
ACCESSION   BQ955883
VERSION     BQ955883.1  GI:22371361
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 935)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2598  row: c  column: 16
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6421959"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
```

EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 39.9%; Score 875; DB 5; Length 935;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        984 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 60

Qy       1044 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 120

Qy       1104 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 180

Qy       1164 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 240

Qy       1224 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 300

Qy       1284 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 360

Qy       1344 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 420

Qy       1404 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 480

Qy       1464 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 540

Qy       1524 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 600

Qy       1584 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 660

Qy       1644 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 720

Qy       1704 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 780

Qy       1764 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 840

Qy       1824 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 1858
              ||||||||||||||||||||||||||||||||||
Db         841 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 875
```

RESULT 10
AL560383
LOCUS AL560383 865 bp mRNA linear EST 02-APR-2004
DEFINITION AL560383 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA clone CS0DG007YD23 5-PRIME, mRNA sequence.
ACCESSION AL560383
VERSION AL560383.3 GI:46185754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31284514.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 6328.r
For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0DG007CB12QP1&c=6328.r.

FEATURES Location/Qualifiers
source 1. .865
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG007YD23"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 39.5%; Score 865; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy          75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy         135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy         195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy         255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy         315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCAGCAGAGGT 879
              |||||||||||||||||||||||||
Db        841 GGCATTCCTCCTGACCAGCAGAGGT 865
```

RESULT 11
CR615117
LOCUS CR615117 967 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DF013YF22 of Fetal brain of Homo sapiens (human).
ACCESSION CR615117
VERSION CR615117.1 GI:50495924
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 967)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 967)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .967
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF013YF22"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"
ORIGIN

Query Match 39.5%; Score 865; DB 4; Length 967;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1223 GAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAA 1282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAA 60

Qy       1283 TGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTT 1342
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         61 TGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTT 120

Qy       1343 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 1402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 180

Qy       1403 CACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 240

Qy       1463 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 300

Qy       1523 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 360

Qy       1583 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 420

Qy       1643 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 480

Qy       1703 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 540

Qy       1763 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 600

Qy       1823 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 660

Qy       1883 CAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGT 1942
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGT 720

Qy       1943 GGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCC 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCC 780

Qy       2003 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 840

Qy       2063 TGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGT 2122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGT 900

Qy       2123 CTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTG 2182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        901 CTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTG 960

Qy       2183 TTGCATT 2189
              |||||||
Db        961 TTGCATT 967


RESULT 12
BU857118
LOCUS       BU857118                919 bp    mRNA    linear   EST 16-OCT-2002
DEFINITION  AGENCOURT_10480719 NIH_MGC_107 Homo sapiens cDNA clone
            IMAGE:6646922 5', mRNA sequence.
ACCESSION   BU857118
VERSION     BU857118.1  GI:24042108
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 919)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2886  row: i  column: 02
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .919
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6646922"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_107"
                     /note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             39.3%;  Score 861;  DB 5;  Length 919;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  861;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 67

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 127

Qy        748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 187

Qy        808 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 247

Qy        868 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 307

Qy        928 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 367

Qy        988 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 427

Qy       1048 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 487

Qy       1108 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 547

Qy       1168 AAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGA 607

Qy       1228 CCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCA 1287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 CCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCA 667

Qy       1288 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTG 1347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTG 727

Qy       1348 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 GGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 787

Qy       1408 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 TGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 847

Qy       1468 AGACCATCACTCTCGAAGTGG 1488
              |||||||||||||||||||||
Db        848 AGACCATCACTCTCGAAGTGG 868


RESULT 13
CR992464
LOCUS       CR992464                895 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR992464 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017P0416 5',
            mRNA sequence.
ACCESSION   CR992464
VERSION     CR992464.1  GI:68286349
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017P0416.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
```

```
           Heubnerweg 6, D-14059 Berlin, Germany
           Tel: +49 30 32639 100
           Fax: +49 30 32639 111
           www.rzpd.de
           This clone is available from RZPD;
           http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017P0416
           contact RZPD (product- support@rzpd.de) for further information.
           Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017P0416"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match            39.2%;  Score 860;  DB 7;  Length 895;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  860;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCAGCA 874
              ||||||||||||||||||||
Db        841 GGCATTCCTCCTGACCAGCA 860


RESULT 14
AL533035
LOCUS       AL533035               1024 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL533035 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN18 5-PRIME, mRNA sequence.
ACCESSION   AL533035
VERSION     AL533035.3  GI:45707944
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1024)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31070867.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN005DG09QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1024
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN18"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match            39.1%;  Score 858;  DB 1;  Length 1024;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches  958;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        65 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 124

Qy        92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       125 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 184

Qy       152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       185 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 244

Qy       212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       245 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 304

Qy       272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       305 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 364

Qy       332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       365 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 424

Qy       392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       425 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 484

Qy       452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       485 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 544

Qy       512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       545 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 604

Qy       572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       605 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 664

Qy       632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       665 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 724

Qy       692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       725 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 784

Qy       752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       785 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 844

Qy       812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       845 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 904

Qy       872 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       905 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAA 964

Qy       932 CATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       965 MATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1024


RESULT 15
CT004068
LOCUS       CT004068                905 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT004068 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016L066 5',
            mRNA sequence.
ACCESSION   CT004068
VERSION     CT004068.1  GI:68297951
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016L066.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016L066
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .905
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9016L066"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9016"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match            39.1%;  Score 857;  DB 7;  Length 905;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  857;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy        75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy       135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy       195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy       255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy       315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360
```

```
Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCA 871
              |||||||||||||||||
Db        841 GGCATTCCTCCTGACCA 857
```

RESULT 16
CR993134

```
LOCUS       CR993134                886 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR993134 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016B1215 5',
            mRNA sequence.
ACCESSION   CR993134
VERSION     CR993134.1  GI:68287019
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 886)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016B1215.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016B1215
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES    Location/Qualifiers
    source  1. .886
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="RZPDp9016B1215"
            /tissue_type="T-Lymphocytes"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="RZPD no.9016"
            /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
            NotI; vector:
            http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
            ; 1st strand cDNA was prepared from mRNA obtained from
            human T-Lymphocytes with a NotI - oligo(dT) primer [5'
            GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
            Double-stranded cDNA was ligated to SalI adaptors,
            digested with NotI and cloned into the NotI and SalI sites
            of the pQE80LSN_cloned vector"
ORIGIN
```

Query Match 38.8%; Score 851; DB 7; Length 886;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         12 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 60

Qy         72 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 120

Qy        132 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 180

Qy        192 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 240

Qy        252 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 300

Qy        312 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 360

Qy        372 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 420

Qy        432 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 480

Qy        492 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 540

Qy        552 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 600

Qy        612 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 660

Qy        672 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 720
```

```
Qy        732 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 780

Qy        792 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAG 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAG 840

Qy        852 GAAGGCATTCC 862
              ||||||||||||
Db        841 GAAGGCATTCC 851


RESULT 17
CX865290
LOCUS       CX865290                867 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_2_A02.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:30929475 5', mRNA sequence.
ACCESSION   CX865290
VERSION     CX865290.1  GI:58545712
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 867)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM1174  row: a  column: 04
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 867.
FEATURES             Location/Qualifiers
     source          1. .867
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30929475"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN
```

```
  Query Match            38.8%;  Score 850;  DB 8;  Length 867;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  850;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 77

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 137

Qy        856 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCA 197

Qy        916 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 CCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAG 257

Qy        976 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 317

Qy       1036 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 377

Qy       1096 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 1155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACT 437

Qy       1156 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 497

Qy       1216 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCA 557

Qy       1276 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 1335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 617

Qy       1336 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 677

Qy       1396 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 737

Qy       1456 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 1515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 CCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCA 797

Qy       1516 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 1575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 AGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTG 857

Qy       1576 GGAAACAGCT 1585
              ||||||||||
Db        858 GGAAACAGCT 867


RESULT 18
AL535099
LOCUS       AL535099                845 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535099 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF008YC07 5-PRIME, mRNA sequence.
ACCESSION   AL535099
VERSION     AL535099.3  GI:45710980
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 845)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30541538.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF008AB04QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .845
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF008YC07"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match           38.5%;  Score 845;  DB 1;  Length 845;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  845;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 73
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 60

Qy         74 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 120

Qy        134 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 180

Qy        194 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 240

Qy        254 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 300

Qy        314 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 360

Qy        374 CAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGC 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGC 420

Qy        434 CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTAC 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTAC 480

Qy        494 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 540

Qy        554 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 600

Qy        614 CCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCT 673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCT 660

Qy        674 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 720

Qy        734 ACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCAC 780

Qy        794 CCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGA 853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGA 840

Qy        854 AGGCA 858
              |||||
Db        841 AGGCA 845


RESULT 19
CX873614
LOCUS       CX873614               843 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_83_A10.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7488309 5', mRNA sequence.
ACCESSION   CX873614
VERSION     CX873614.1  GI:58556788
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15819  row: b  column: 19
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 843.
FEATURES             Location/Qualifiers
     source          1. .843
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7488309"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
```

/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
RNA obtained from human embryonic stem cells isolated from
the inner cell mass of blastocyst stage embryos and
differentiated to an early neural progenitor cell type.
Cell line id and NIH Registry designation is BG01.
Positive for Nestin and Musashi expression. Passage number
18. cDNA primed using oligo-dT primer:
5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. This primary library
is non-normalized (normalized primary library is
NIH_MGC_259). It was constructed by Express Genomics
(Frederick, MD). Sequence ends have been trimmed to
exclude vector and regions below Phred quality 16. Note:
this is a Mammalian Gene Collection library."

ORIGIN

Query Match 38.5%; Score 843; DB 8; Length 843;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        807 CCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCT 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCT 60

Qy        867 GACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGAC 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGAC 120

Qy        927 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 180

Qy        987 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACC 1046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACC 240

Qy       1047 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 300

Qy       1107 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 1166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 360

Qy       1167 AAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 1226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 420

Qy       1227 ACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTC 1286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTC 480

Qy       1287 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCT 1346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCT 540

Qy       1347 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 600

Qy       1407 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 660

Qy       1467 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 720

Qy       1527 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 1586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 780

Qy       1587 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 840

Qy       1647 CTC 1649
              |||
Db        841 CTC 843
```

RESULT 20
CR997691
LOCUS CR997691 881 bp mRNA linear EST 28-JUN-2005
DEFINITION CR997691 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017E119 5', mRNA sequence.
ACCESSION CR997691
VERSION CR997691.1 GI:68291576
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9017E119.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017E119
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9017E119"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 38.4%; Score 842; DB 7; Length 881;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 60

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 120

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 180

Qy        504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 240

Qy        564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 300

Qy        624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 360

Qy        684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 420

Qy        744 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 480

Qy        804 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 540

Qy        864 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 600

Qy        924 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 660

Qy        984 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 720

Qy       1044 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 780

Qy       1104 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 840

Qy       1164 CA 1165
              ||
Db          841 CA 842
```

```
RESULT 21
AL560428
LOCUS       AL560428                937 bp    mRNA    linear   EST 02-APR-2004
DEFINITION  AL560428 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG006YE05 5-PRIME, mRNA sequence.
ACCESSION   AL560428
VERSION     AL560428.3  GI:46185799
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 937)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31284559.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DG006AC03QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .937
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG006YE05"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN
```

```
  Query Match             38.4%;  Score 841;  DB 1;  Length 937;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  841;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 G 856
                |
Db        841 G 841


RESULT 22
BX419560
LOCUS       BX419560                847 bp    mRNA    linear   EST 01-MAY-2004
DEFINITION  BX419560 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF018YM19 5-PRIME, mRNA sequence.
ACCESSION   BX419560
VERSION     BX419560.2  GI:46925988
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30642520.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF018AG10QP1&c=6328.r.
FEATURES              Location/Qualifiers
     source           1. .847
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DF018YM19"
                      /tissue_type="FETAL BRAIN"
                      /dev_stage="fetal"
                      /clone_lib="Homo sapiens FETAL BRAIN"
                      /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                      was primed with a NotI-oligo(dT) primer. Five prime end
                      enriched, double-strand cDNA was digested with Not I and
                      cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                      vector. Library was not normalized."
ORIGIN

  Query Match           38.3%;  Score 840;  DB 5;  Length 847;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  840;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 67

Qy         81 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 127

Qy        141 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 187

Qy        201 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 247

Qy        261 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 307

Qy        321 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 367

Qy        381 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 427

Qy        441 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 487

Qy        501 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 547

Qy        561 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 607

Qy        621 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 667

Qy        681 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 727

Qy        741 CTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAG 800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAG 787

Qy        801 GTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATT 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        788 GTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATT 847


RESULT 23
CR981232
LOCUS       CR981232                915 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR981232 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017B1419 5',
            mRNA sequence.
ACCESSION   CR981232
VERSION     CR981232.1  GI:68219566
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
```

```
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017B1419.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017B1419
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .915
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9017B1419"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           38.1%;  Score 836;  DB 7;  Length 915;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  836;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 60

Qy         72 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 120

Qy        132 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 180

Qy        192 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 240

Qy        252 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 300

Qy        312 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 360

Qy        372 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 420

Qy        432 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 480

Qy        492 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 540

Qy        552 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 600

Qy        612 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 660

Qy        672 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 720

Qy        732 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 780

Qy        792 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 847
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 836


RESULT 24
BQ073574/c
LOCUS       BQ073574               1026 bp    mRNA    linear   EST 02-APR-2002
DEFINITION  AGENCOURT_7046537 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806408
            5', mRNA sequence.
ACCESSION   BQ073574
VERSION     BQ073574.1  GI:19902620
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1026)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2047  row: c  column: 17
            High quality sequence stop: 667.
FEATURES              Location/Qualifiers
     source           1. .1026
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5806408"
                      /tissue_type="epidermoid carcinoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_101"
                      /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
                      XhoI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Library constructed by Ling Hong in the
                      laboratory of Gerald M. Rubin (University of California,
                      Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                      Superscript II RT (Life Technologies). Note: this is a
```

```
                    NIH_MGC Library."
ORIGIN

  Query Match             38.0%;  Score 833;  DB 3;  Length 1026;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  833;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1360 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 1419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 AAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGC 792

Qy       1420 TCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         791 TCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 732

Qy       1480 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         731 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 672

Qy       1540 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         671 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 612

Qy       1600 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAG 1659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         611 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAG 552

Qy       1660 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGC 1719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGC 492

Qy       1720 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         491 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 432

Qy       1780 ACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         431 ACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 372

Qy       1840 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 1899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         371 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 312

Qy       1900 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCA 1959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         311 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCA 252

Qy       1960 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 2019
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         251 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGT 192

Qy       2020 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 2079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         191 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 132

Qy       2080 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 2139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         131 AAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTA 72

Qy       2140 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          71 AGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 19


RESULT 25
AL535209
LOCUS       AL535209                1104 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535209 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF008YB24 5-PRIME, mRNA sequence.
ACCESSION   AL535209
VERSION     AL535209.3  GI:45711087
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1104)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30541755.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF008DA12QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1104
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF008YB24"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             38.0%;  Score 832;  DB 1;  Length 1104;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  882;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 77

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 137

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 197

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 257

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 317

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        318 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 377

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 437

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 497
```

```
Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 557

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 617

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 677

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 737

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 797

Qy        812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCA 857

Qy        872 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 914
              |||||||||||||||||||||||||||||||||||||||||||
Db        858 GCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 900
```

RESULT 26
CR998818
LOCUS CR998818 913 bp mRNA linear EST 28-JUN-2005
DEFINITION CR998818 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017M0415 5', mRNA sequence.
ACCESSION CR998818
VERSION CR998818.1 GI:68292703
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 913)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9017M0415.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017M0415
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .913
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9017M0415"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"
ORIGIN

Query Match 37.9%; Score 830; DB 7; Length 913;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 830
```

```
RESULT 27
CX871103
LOCUS       CX871103                829 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_53_C01.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7485675 5', mRNA sequence.
ACCESSION   CX871103
VERSION     CX871103.1  GI:58554277
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 829)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15812  row: e  column: 01
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 829.
FEATURES             Location/Qualifiers
     source          1. .829
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7485675"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             37.8%;  Score 829;  DB 8;  Length 829;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  829;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 60

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 120

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 180

Qy        909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 240

Qy        969 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAA 300

Qy       1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 360

Qy       1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 420

Qy       1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 480

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 540

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 600

Qy       1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 660

Qy       1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 720

Qy       1449 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 780

Qy       1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 829


RESULT 28
CR996018
LOCUS       CR996018                839 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR996018 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017J0713 5',
            mRNA sequence.
ACCESSION   CR996018
VERSION     CR996018.1  GI:68289903
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 839)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017J0713.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
```

```
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017J0713
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .839
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9017J0713"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             37.7%;  Score 827;  DB 7;  Length 839;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  827;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1251 CTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1310
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 60

Qy       1311 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1370
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 120

Qy       1371 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGA 1430
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGA 180

Qy       1431 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 1490
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 240

Qy       1491 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 300

Qy       1551 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 360

Qy       1611 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1670
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 420

Qy       1671 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1730
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 480

Qy       1731 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1790
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 540

Qy       1791 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1850
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 600

Qy       1851 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 1910
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 660

Qy       1911 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 1970
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 720

Qy       1971 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCT 2030
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCT 780

Qy       2031 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 2077
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCA 827


RESULT 29
CX866263
LOCUS       CX866263                 831 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_9_E11.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7471967 5', mRNA sequence.
ACCESSION   CX866263
VERSION     CX866263.1  GI:58547834
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: HESC4_9_E11.b1_A037
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15776  row: i  column: 21
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 831.
FEATURES              Location/Qualifiers
     source           1. .831
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:7471967"
                      /sex="male"
                      /tissue_type="embryonic stem"
                      /cell_type="human embryonic stem cells"
                      /cell_line="BG01"
                      /lab_host="DH10B-T1 phage-resistant E. coli"
                      /clone_lib="NIH_MGC_262"
                      /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                      RNA obtained from human embryonic stem cells isolated from
                      the inner cell mass of blastocyst stage embryos and
                      differentiated to an early neural progenitor cell type.
                      Cell line id and NIH Registry designation is BG01.
                      Positive for Nestin and Musashi expression. Passage number
                      18. cDNA primed using oligo-dT primer:
```

5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 37.6%; Score 825; DB 8; Length 831;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 66

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 126

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 186

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 246

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 306

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 366

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 426

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 486

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 546

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 606

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 666

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 726

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 786

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAG 839
              |||||||||||||||||||||||||||||||||||||||||||||
Db        787 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAG 831
```

RESULT 30
CR990940
LOCUS CR990940 824 bp mRNA linear EST 28-JUN-2005

DEFINITION CR990940 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016B014 5', mRNA sequence.
ACCESSION CR990940
VERSION CR990940.1 GI:68284825
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016B014.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016B014
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016B014"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 37.6%; Score 824; DB 7; Length 824;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 824; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 60

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 120

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 180

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 240
```

```
Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 300

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 360

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 420

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 480

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 540

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 600

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 660

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 720

Qy        855 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 780

Qy        915 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 958
              |||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 824
```

RESULT 31
CD517425

```
LOCUS       CD517425                902 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14375655 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30395903 5', mRNA sequence.
ACCESSION   CD517425
VERSION     CD517425.1  GI:31449143
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 902)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM476  row: l  column: 24
            High quality sequence stop: 649.
FEATURES              Location/Qualifiers
     source           1. .902
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:30395903"
                      /tissue_type="White Matter"
                      /dev_stage="Unknown"
                      /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                      /clone_lib="NIH_MGC_181"
                      /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                      (destroyed); Library is oligo-dT primed and directionally
                      cloned (EcoRV site is destroyed upon cloning). Average
                      insert size 1.42 kb. Library was constructed by
                      (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match             37.4%;  Score 820;  DB 6;  Length 902;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  820;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1175 CACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 60

Qy       1235 TGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 120

Qy       1295 GATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 180

Qy       1355 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 240

Qy       1415 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 300

Qy       1475 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 360

Qy       1535 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 420

Qy       1595 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 480

Qy       1655 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 1714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 540

Qy       1715 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 600

Qy       1775 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 660

Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 720

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 780

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 1994
```

```
                  ||||||||||||||||||||||||||||||||||||||||
Db        781 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 820


RESULT 32
BM541766
LOCUS       BM541766                1086 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6439070 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531666
            5', mRNA sequence.
ACCESSION   BM541766
VERSION     BM541766.1  GI:18770693
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1086)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12213  row: l  column: 03
            High quality sequence stop: 720.
FEATURES             Location/Qualifiers
     source          1. .1086
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5531666"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match             37.4%;  Score 819;  DB 3;  Length 1086;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  819;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          25 GGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGA 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGA 60

Qy          85 AGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATG 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATG 120

Qy         145 TCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTG 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTG 180

Qy         205 CTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 240

Qy         265 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 324
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 300

Qy         325 GCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGA 384
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGA 360

Qy         385 TCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 444
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 420

Qy         445 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 504
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 480

Qy         505 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 564
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 540

Qy         565 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 624
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 600

Qy         625 AAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 684
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 660

Qy         685 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 744
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 720

Qy         745 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 804
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 780

Qy         805 AGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
               |||||||||||||||||||||||||||||||||||||||
Db         781 AGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 819


RESULT 33
BM450657
LOCUS       BM450657                1009 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6394477 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494351
            5', mRNA sequence.
ACCESSION   BM450657
VERSION     BM450657.1  GI:18499697
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1009)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12119  row: i  column: 08
            High quality sequence stop: 667.
FEATURES             Location/Qualifiers
     source          1. .1009
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5494351"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
```

Average insert size 1.75 kb. Library constructed by Life Technologies."
ORIGIN

Query Match 37.3%; Score 818; DB 3; Length 1009;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy      75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy     135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy     195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy     255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy     315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy     375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy     435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy     495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy     555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy     615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy     675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy     735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy     795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 832
           ||||||||||||||||||||||||||||||||||||||
Db     781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 818
```

RESULT 34
BQ439334
LOCUS BQ439334 885 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7761806 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6018275 5', mRNA sequence.
ACCESSION BQ439334
VERSION BQ439334.1 GI:21178410
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13218 row: o column: 12
High quality sequence stop: 715.
FEATURES Location/Qualifiers
source 1. .885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6018275"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
ORIGIN

Query Match 37.1%; Score 814; DB 5; Length 885;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 814; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     490 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 60

Qy     550 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 120

Qy     610 AGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAAC 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAAC 180

Qy     670 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 240

Qy     730 TGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCA 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCA 300

Qy     790 TCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACA 849
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACA 360

Qy     850 AGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATG 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATG 420

Qy     910 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTC 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTC 480

Qy     970 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAG 1029
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAG 540

Qy       1030 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 600

Qy       1090 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGT 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGT 660

Qy       1150 CTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 720

Qy       1210 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTG 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTG 780

Qy       1270 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1303
              ||||||||||||||||||||||||||||||||||
Db        781 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 814


RESULT 35
AL541568
LOCUS       AL541568                939 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL541568 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE003YA21
            5-PRIME, mRNA sequence.
ACCESSION   AL541568
VERSION     AL541568.3  GI:45717155
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:30545869.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DE003AA11QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE003YA21"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match            37.1%;  Score 814;  DB 1;  Length 939;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  864;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 95

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 155

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 215

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 275

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        276 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 335

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 395

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 455

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 515

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 575

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 635

Qy        649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 695

Qy        709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 755

Qy        769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        756 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 815

Qy        829 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        816 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 875

Qy        889 CCGGAAAGCAGCTGGAAGATGGGCG 913
              |||||||||||||||||||||||||
Db        876 CCGGAAAGCAGCTGGAAGATGGGCG 900


RESULT 36
BU153047
LOCUS       BU153047                897 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_7945088 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6010685
            5', mRNA sequence.
ACCESSION   BU153047
VERSION     BU153047.1  GI:22666579
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 897)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13199  row: c  column: 06
            High quality sequence stop: 725.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6010685"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match           37.1%;  Score 813;  DB 5;  Length 897;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  813;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         26 GTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAA 85
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAA 60

Qy         86 GACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGT 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGT 120

Qy        146 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGC 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGC 180

Qy        206 TGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 240

Qy        266 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 300

Qy        326 CAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGAT 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGAT 360

Qy        386 CCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCT 420

Qy        446 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGT 480

Qy        506 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 540

Qy        566 CCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGA 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGA 600

Qy        626 AGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCG 660

Qy        686 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAG 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAG 720

Qy        746 AGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGA 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGA 780

Qy        806 GCCCAGTGACACTATCGAGAACGTCAAAGCAAA 838
              ||||||||||||||||||||||||||||||||
Db        781 GCCCAGTGACACTATCGAGAACGTCAAAGCAAA 813


RESULT 37
BX458876
LOCUS       BX458876               1009 bp    mRNA    linear   EST 06-MAY-2004
DEFINITION  BX458876 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010YM03
            5-PRIME, mRNA sequence.
ACCESSION   BX458876
VERSION     BX458876.2  GI:47069749
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1009)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 22, 2003 this sequence version replaced gi:31033016.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1009
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE010YM03"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match           37.0%;  Score 812;  DB 5;  Length 1009;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  862;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 83

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 143

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 203

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 263

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        264 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 323

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 383

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 443

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 503

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 563

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 623

Qy        649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 683

Qy        709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 743

Qy        769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        744 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 803

Qy        829 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 888
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        804 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 863

Qy        889 CCGGAAAGCAGCTGGAAGATGGG 911
                |||||||||||||||||||||||
Db        864 CCGGAAAGCAGCTGGAAGATGGG 886


RESULT 38
CR990631
LOCUS       CR990631                831 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR990631 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017C1312 5',
            mRNA sequence.
ACCESSION   CR990631
VERSION     CR990631.1  GI:68284516
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017C1312.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017C1312
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017C1312"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             37.0%;  Score 811;  DB 7;  Length 831;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  811;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1302 GACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 1361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 60

Qy       1362 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 1421
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 120

Qy       1422 CGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 1481
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 180

Qy       1482 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1541
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 240

Qy       1542 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 1601
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 300

Qy       1602 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGT 1661
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGT 360

Qy       1662 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 1721
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 420

Qy       1722 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1781
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 480

Qy       1782 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1841
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 540

Qy       1842 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 600

Qy       1902 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 1961
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATC 660

Qy       1962 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 2021
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 720

Qy       2022 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 2081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 780

Qy       2082 GAGTCCACTCTGCACTTGGTCCTGCGCTTGA 2112
             ||||||||||||||||||||||||||||||||
Db        781 GAGTCCACTCTGCACTTGGTCCTGCGCTTGA 811


RESULT 39
CX865249
LOCUS       CX865249                819 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_1_F12.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:30929734 5', mRNA sequence.
ACCESSION   CX865249
VERSION     CX865249.1  GI:58545617
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM1174  row: k  column: 23
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 819.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30929734"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             37.0%;  Score 810;  DB 8;  Length 819;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  810;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        612 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 69

Qy        672 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 129

Qy        732 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 189

Qy        792 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAG 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAG 249

Qy        852 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 309

Qy        912 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 369

Qy        972 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTG 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTG 429

Qy       1032 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCT 489

Qy       1092 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCT 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCT 549

Qy       1152 GACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 GACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 609

Qy       1212 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGAC 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGAC 669

Qy       1272 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAG 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAG 729

Qy       1332 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 AGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATC 789

Qy       1392 CAGAAAGAGTCCACCCTGCACCTGGTGCTC 1421
              |||||||||||||||||||||||||||||
Db        790 CAGAAAGAGTCCACCCTGCACCTGGTGCTC 819


RESULT 40
```

```
BM928821
LOCUS       BM928821                1036 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6683065 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:5771435
            5', mRNA sequence.
ACCESSION   BM928821
VERSION     BM928821.1  GI:19379200
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1036)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12836  row: b  column: 12
            High quality sequence stop: 643.
FEATURES             Location/Qualifiers
     source          1. .1036
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5771435"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match             37.0%;  Score 810;  DB 3;  Length 1036;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  810;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 60

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 120

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 180

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 240

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 300

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 360

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 420

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 480

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 540

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 600

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 660

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 720

Qy        742 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGG 780

Qy        802 TCGAGCCCAGTGACACTATCGAGAACGTCA 831
              |||||||||||||||||||||||||||||
Db        781 TCGAGCCCAGTGACACTATCGAGAACGTCA 810


RESULT 41
CR999015
LOCUS       CR999015                898 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR999015 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017G0318 5',
            mRNA sequence.
ACCESSION   CR999015
VERSION     CR999015.1  GI:68292900
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 898)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017G0318.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017G0318
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .898
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017G0318"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
```

/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 36.9%; Score 809; DB 7; Length 898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy      75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy     135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy     195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy     255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy     315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy     375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy     435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy     495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy     555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy     615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy     675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy     735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy     795 CTTGAGGTCGAGCCCAGTGACACTATCGA 823
           |||||||||||||||||||||||||||||
Db     781 CTTGAGGTCGAGCCCAGTGACACTATCGA 809
```

RESULT 42
CX163519
LOCUS CX163519 824 bp mRNA linear EST 23-DEC-2004
DEFINITION HESC2_5_H05.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone IMAGE:7467107 5', mRNA sequence.
ACCESSION CX163519
VERSION CX163519.1 GI:56793599
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM15763 row: o column: 09
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 824.

FEATURES Location/Qualifiers
source 1. .824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7467107"
/sex="male"
/tissue_type="human embryonic stem cells differentiated to an early endodermal cell type"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_258"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early endodermal cell type. Cell line id and NIH Registry designation is BG01. Positive for GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP expression. Passage number 40. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 36.8%; Score 807; DB 8; Length 824;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      41 TGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAA 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18 TGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAA 77
```

```
Qy        101 GACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 137

Qy        161 AGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGA 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGA 197

Qy        221 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCT 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCT 257

Qy        281 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 317

Qy        341 TGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGG 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 TGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGG 377

Qy        401 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCAC 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCAC 437

Qy        461 CCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGG 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 CCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGG 497

Qy        521 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCC 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCC 557

Qy        581 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGA 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGA 617

Qy        641 TCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTA 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 TCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTA 677

Qy        701 CAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAAT 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 CAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAAT 737

Qy        761 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTAT 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTAT 797

Qy        821 CGAGAACGTCAAAGCAAAGATCCAAGA 847
              ||||||||||||||||||||||||||||
Db        798 CGAGAACGTCAAAGCAAAGATCCAAGA 824


RESULT 43
BM551807
LOCUS       BM551807                 895 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6544760 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5548187
            5', mRNA sequence.
ACCESSION   BM551807
VERSION     BM551807.1  GI:18789144
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12256  row: 1  column: 12
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5548187"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             36.8%;  Score 807;  DB 3;  Length 895;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  807;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1313 CATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCAC 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCAC 60

Qy       1373 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGG 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGG 120

Qy       1433 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCC 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCC 180

Qy       1493 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 240

Qy       1553 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 300

Qy       1613 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 360

Qy       1673 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 420

Qy       1733 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 480

Qy       1793 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 540

Qy       1853 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 600

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 660

Qy       1973 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 2032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        661 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGG 720

Qy       2033 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 2092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCT 780

Qy       2093 GCACTTGGTCCTGCGCTTGAGGGGGGG 2119
              ||||||||||||||||||||||||||||
Db        781 GCACTTGGTCCTGCGCTTGAGGGGGGG 807


RESULT 44
CD244749
LOCUS       CD244749                889 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14097437 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30376725 5', mRNA sequence.
ACCESSION   CD244749
VERSION     CD244749.1  GI:31005213
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 889)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM439  row: m  column: 22
            High quality sequence stop: 685.
FEATURES             Location/Qualifiers
     source          1. .889
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30376725"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            36.6%;  Score 803;  DB 6;  Length 889;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  803;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 60

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 240

Qy        996 AAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 1055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAAT 300

Qy       1056 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTT 360

Qy       1116 GCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCC 420

Qy       1176 ACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 480

Qy       1236 GGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 540

Qy       1296 ATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAG 600

Qy       1356 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 660

Qy       1416 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 720

Qy       1476 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 780

Qy       1536 GAAGGCATCCCTCCTGACCAGCA 1558
              |||||||||||||||||||||||
Db        781 GAAGGCATCCCTCCTGACCAGCA 803


RESULT 45
BU187089
LOCUS       BU187089                919 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7780299 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137876
            5', mRNA sequence.
ACCESSION   BU187089
VERSION     BU187089.1  GI:22701073
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 919)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13452  row: f  column: 21
            High quality sequence stop: 594.
FEATURES             Location/Qualifiers
```

```
     source          1. .919
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6137876"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             36.6%;  Score 803;  DB 5;  Length 919;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  803;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        210 AAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 60

Qy        270 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 120

Qy        330 ACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAG 180

Qy        390 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 240

Qy        450 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 300

Qy        510 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 360

Qy        570 GAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGC 420

Qy        630 ATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 480

Qy        690 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGT 749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGT 540

Qy        750 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCC 600

Qy        810 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGAC 660

Qy        870 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 720

Qy        930 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 780

Qy        990 TTCGTGAAGACCCTGACTGGTAA 1012
              |||||||||||||||||||||||
Db        781 TTCGTGAAGACCCTGACTGGTAA 803
```

```
RESULT 46
DR762064
LOCUS       DR762064                801 bp    mRNA    linear   EST 21-JUL-2005
DEFINITION  HESC4_136_A07.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7971472 5', mRNA sequence.
ACCESSION   DR762064
VERSION     DR762064.1  GI:71054764
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM 17088  row: b  column: 14
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 801.
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7971472"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             36.5%;  Score 801;  DB 8;  Length 801;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  801;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120
Qy       136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
Db       121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180
Qy       196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
Db       181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240
Qy       256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
Db       241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300
Qy       316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
Db       301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360
Qy       376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
Db       361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420
Qy       436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
Db       421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480
Qy       496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
Db       481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540
Qy       556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
Db       541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600
Qy       616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
Db       601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660
Qy       676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
Db       661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720
Qy       736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
Db       721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780
Qy       796 TTGAGGTCGAGCCCAGTGACA 816
Db       781 TTGAGGTCGAGCCCAGTGACA 801
```

RESULT 47
AL555014
LOCUS AL555014 1022 bp mRNA linear EST 30-MAR-2004
DEFINITION AL555014 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens cDNA clone CS0DK009YF01 5-PRIME, mRNA sequence.
ACCESSION AL555014
VERSION AL555014.3 GI:45859758
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31276823.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6328.r
For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0DK009CC01QP1&c=6328.r.
FEATURES Location/Qualifiers
source 1. .1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK009YF01"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match 36.5%; Score 801; DB 1; Length 1022;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
Db         8 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 67
Qy        80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
Db        68 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 127
Qy       140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
Db       128 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 187
Qy       200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
Db       188 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 247
Qy       260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
Db       248 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 307
Qy       320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
Db       308 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 367
Qy       380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
Db       368 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 427
Qy       440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
Db       428 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 487
Qy       500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
Db       488 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 547
Qy       560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
Db       548 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 607
Qy       620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
Db       608 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 667
Qy       680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
```

```
Db        668 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 727

Qy        740 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 787

Qy        800 GGTCGAGCCCAGTGACACTAT 820
              |||||||||||||||||||||
Db        788 GGTCGAGCCCAGTGACACTAT 808


RESULT 48
AL537821
LOCUS       AL537821                1042 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL537821 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF028YH11 5-PRIME, mRNA sequence.
ACCESSION   AL537821
VERSION     AL537821.3  GI:45713563
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1042)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31262422.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF028CD06QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1042
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF028YH11"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             36.5%;  Score 801;  DB 1;  Length 1042;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGWCCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAG 840

Qy        856 GCATTCCTCCTG 867
              ||||||||||||
Db        841 GCATTCCTCCTG 852


RESULT 49
AL560244
LOCUS       AL560244                996 bp    mRNA    linear   EST 02-APR-2004
DEFINITION  AL560244 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG002YC21 5-PRIME, mRNA sequence.
ACCESSION   AL560244
VERSION     AL560244.3  GI:46185615
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 996)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31284375.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
```

```
           into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
           was not normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
           This sequence belongs to sequence cluster 6328.r
           For more information about this cluster, see
           http://www.genoscope.cns.fr/cdna?s=CS0DG002AB11QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .996
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DG002YC21"
                     /tissue_type="B CELLS (RAMOS CELL LINE)"
                     /cell_line="RAMOS CELL LINE"
                     /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match           36.5%;  Score 800;  DB 1;  Length 996;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  800;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGA 814
              ||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGA 800


RESULT 50
CR984609
LOCUS       CR984609                853 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR984609 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017L2421 5',
            mRNA sequence.
ACCESSION   CR984609
VERSION     CR984609.1  GI:68222943
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 853)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017L2421.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017L2421
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .853
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017L2421"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           36.5%;  Score 799;  DB 7;  Length 853;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  849;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        348 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCT 63
```

```
Qy        408 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 123

Qy        468 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 183

Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 243

Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 303

Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 363

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 423

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 483

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 543

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 603

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 663

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        664 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 723

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 783

Qy       1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 843

Qy       1188 GTGCTCCGTC 1197
              ||||||||||
Db        844 GTGCTCCGTC 853


RESULT 51
CR994031
LOCUS       CR994031                859 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR994031 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017D2218 5',
            mRNA sequence.
ACCESSION   CR994031
VERSION     CR994031.1  GI:68287916
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017D2218.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017D2218
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .859
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017D2218"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN


  Query Match             36.4%;  Score 797;  DB 7;  Length 859;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        806 GCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCC 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCC 63

Qy        866 TGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGA 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 TGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGA 123

Qy        926 CTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 183

Qy        986 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACAC 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACAC 243

Qy       1046 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 CATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAG 303

Qy       1106 GTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCA 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCA 363

Qy       1166 GAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 423
```

```
Qy       1226 GACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGT 1285
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 GACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGT 483

Qy       1286 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 1345
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 CAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGC 543

Qy       1346 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 1405
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 TGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCAC 603

Qy       1406 CCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 663

Qy       1466 TAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGAT 1525
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 TAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGAT 723

Qy       1526 CCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 1585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 CCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 783

Qy       1586 GGAAGATGGACGCACCC 1602
             |||||||||||||||||
Db        784 GGAAGATGGACGCACCC 800
```

RESULT 52
AL535900

```
LOCUS       AL535900                965 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535900 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF013YF22 5-PRIME, mRNA sequence.
ACCESSION   AL535900
VERSION     AL535900.3  GI:45711748
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30542878.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF013DC11QP1&c=6328.r.
FEATURES                Location/Qualifiers
     source             1. .965
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CS0DF013YF22"
                        /tissue_type="FETAL BRAIN"
                        /dev_stage="fetal"
                        /clone_lib="Homo sapiens FETAL BRAIN"
                        /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                        was primed with a NotI-oligo(dT) primer. Five prime end
                        enriched, double-strand cDNA was digested with Not I and
                        cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                        vector. Library was not normalized."
```

ORIGIN

```
  Query Match             36.4%;  Score 797;  DB 1;  Length 965;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  897;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1235 TGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1294
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          13 TGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 72

Qy       1295 GATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACA 1354
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          73 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 132

Qy       1355 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         133 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 192

Qy       1415 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         193 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 252

Qy       1475 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         253 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 312

Qy       1535 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         313 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 372

Qy       1595 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         373 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 432

Qy       1655 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 1714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         433 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 492

Qy       1715 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         493 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 552

Qy       1775 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         553 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 612

Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         613 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 672

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         673 GCAAATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 732

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         733 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 792

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         793 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 852

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCC 2133
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         853 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCC 911
```

RESULT 53
BU194577

```
LOCUS       BU194577                876 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7970058 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6147303
            5', mRNA sequence.
ACCESSION   BU194577
```

```
VERSION     BU194577.1  GI:22708561
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 876)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13476  row: o  column: 16
            High quality sequence stop: 602.
FEATURES             Location/Qualifiers
     source          1. .876
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6147303"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             36.3%;  Score 796;  DB 5;  Length 876;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  796;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 62

Qy         89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 122

Qy        149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 182

Qy        209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 242

Qy        269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 302

Qy        329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 362

Qy        389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 422

Qy        449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 482

Qy        509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 542

Qy        569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 602

Qy        629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 662

Qy        689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 722

Qy        749 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCC 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCC 782

Qy        809 CAGTGACACTATCGAG 824
              ||||||||||||||||
Db        783 CAGTGACACTATCGAG 798


RESULT 54
CR998795
LOCUS       CR998795                824 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR998795 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017O1638 5',
            mRNA sequence.
ACCESSION   CR998795
VERSION     CR998795.1  GI:68292680
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 824)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017O1638.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017O1638
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .824
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017O1638"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
```

; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 36.3%; Score 795; DB 7; Length 824;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        882 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 60

Qy        942 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1001
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 120

Qy       1002 CTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 180

Qy       1062 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 1121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 240

Qy       1122 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 1181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 300

Qy       1182 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 360

Qy       1242 ACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 1301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 420

Qy       1302 GACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 1361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 480

Qy       1362 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 1421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 540

Qy       1422 CGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 1481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 600

Qy       1482 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 660

Qy       1542 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 1601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACC 720

Qy       1602 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGT 1661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGT 780

Qy       1662 GGGATGCAGATCTTC 1676
              |||||||||||||||
Db        781 GGGATGCAGATCTTC 795
```

RESULT 55
BQ225402
LOCUS BQ225402 873 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7546060 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6025106.
5', mRNA sequence.
ACCESSION BQ225402
VERSION BQ225402.1 GI:20406802
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 873)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13236 row: l column: 03
High quality sequence stop: 606.
FEATURES Location/Qualifiers
source 1. .873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6025106"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
ORIGIN

Query Match 36.3%; Score 795; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 60

Qy         89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 120

Qy        149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 180

Qy        209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 240

Qy        269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 300

Qy        329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 360

Qy        389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 420

Qy        449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 480

Qy        509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 540

Qy        569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 600

Qy        629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 660

Qy        689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 720

Qy        749 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCC 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCC 780

Qy        809 CAGTGACACTATCGA 823
              |||||||||||||||
Db        781 CAGTGACACTATCGA 795


RESULT 56
CF552717
LOCUS       CF552717                879 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15595429 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529113 5', mRNA sequence.
ACCESSION   CF552717
VERSION     CF552717.1  GI:34889551
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM616  row: c  column: 10
            High quality sequence start: 28
            High quality sequence stop: 674.
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529113"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match           36.3%;  Score 795;  DB 6;  Length 879;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  795;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1306 AGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 1365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 AGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 98

Qy       1366 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 158

Qy       1426 TTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAG 1485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 TTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAG 218

Qy       1486 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 278

Qy       1546 CTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 338

Qy       1606 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 1665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 398

Qy       1666 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTG 1725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTG 458

Qy       1726 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 1785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGC 518

Qy       1786 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 1845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 578

Qy       1846 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 1905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 638

Qy       1906 TGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGA 1965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 TGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGA 698

Qy       1966 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCT 2025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCT 758

Qy       2026 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 2085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGT 818

Qy       2086 CCACTCTGCACTTGG 2100
              |||||||||||||||
Db        819 CCACTCTGCACTTGG 833


RESULT 57
CX866677
LOCUS       CX866677                794 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_12_D03.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7471928 5', mRNA sequence.
ACCESSION   CX866677
VERSION     CX866677.1  GI:58548721
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: HESC4_12_D03.b1_A037
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15776  row: h  column: 06
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 794.
FEATURES              Location/Qualifiers
     source           1. .794
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:7471928"
                      /sex="male"
                      /tissue_type="embryonic stem"
                      /cell_type="human embryonic stem cells"
                      /cell_line="BG01"
                      /lab_host="DH10B-T1 phage-resistant E. coli"
                      /clone_lib="NIH_MGC_262"
                      /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                      RNA obtained from human embryonic stem cells isolated from
                      the inner cell mass of blastocyst stage embryos and
                      differentiated to an early neural progenitor cell type.
                      Cell line id and NIH Registry designation is BG01.
                      Positive for Nestin and Musashi expression. Passage number
                      18. cDNA primed using oligo-dT primer:
                      5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                      the EcoRV/NotI sites of pExpress-1. This primary library
                      is non-normalized (normalized primary library is
                      NIH_MGC_259). It was constructed by Express Genomics
                      (Frederick, MD). Sequence ends have been trimmed to
                      exclude vector and regions below Phred quality 16. Note:
                      this is a Mammalian Gene Collection library."
ORIGIN

  Query Match           36.2%;  Score 794;  DB 8;  Length 794;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  794;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          21 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 60

Qy          81 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 120

Qy         141 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 180

Qy         201 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 240

Qy         261 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 300

Qy         321 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 360

Qy         381 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 420

Qy         441 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 480

Qy         501 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 540

Qy         561 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 600

Qy         621 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 660

Qy         681 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 720

Qy         741 CTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAG 800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAG 780

Qy         801 GTCGAGCCCAGTGA 814
               ||||||||||||||
Db         781 GTCGAGCCCAGTGA 794


RESULT 58
BQ889352
LOCUS       BQ889352                945 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_7981088 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6185674 5', mRNA sequence.
ACCESSION   BQ889352
VERSION     BQ889352.1  GI:22281366
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13576  row: n  column: 11
            High quality sequence start: 21
            High quality sequence stop: 664.
FEATURES              Location/Qualifiers
     source           1. .945
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
```

/clone="IMAGE:6185674"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 36.2%; Score 794; DB 5; Length 945;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 90

Qy        185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 150

Qy        245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 210

Qy        305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 270

Qy        365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 330

Qy        425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 390

Qy        485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 450

Qy        545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 510

Qy        605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 570

Qy        665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 630

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 690

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCA 750

Qy        845 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 AGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 810

Qy        905 AGATGGGCGCACCC 918
              ||||||||||||||
Db        811 AGATGGGCGCACCC 824
```

RESULT 59
BM472268
LOCUS BM472268 963 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6469625 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5576452 5', mRNA sequence.
ACCESSION BM472268
VERSION BM472268.1 GI:18521310
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12328 row: f column: 05
High quality sequence stop: 715.

FEATURES Location/Qualifiers
source 1. .963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5576452"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 36.1%; Score 792; DB 3; Length 963;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1401 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 60

Qy       1461 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 120

Qy       1521 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 180

Qy       1581 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 240

Qy       1641 CTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 300
```

```
Qy       1701 ATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 1760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 360

Qy       1761 AAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 1820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGAT 420

Qy       1821 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGT 1880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGT 480

Qy       1881 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 1940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 540

Qy       1941 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATC 2000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATC 600

Qy       2001 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 2060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 660

Qy       2061 TCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGT 2120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGT 720

Qy       2121 GTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGT 2180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGT 780

Qy       2181 TGTTGCATTCCC 2192
             ||||||||||||
Db        781 TGTTGCATTCCC 792
```

RESULT 60
CR996412
LOCUS CR996412 842 bp mRNA linear EST 28-JUN-2005
DEFINITION CR996412 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017E213 5', mRNA sequence.
ACCESSION CR996412
VERSION CR996412.1 GI:68290297
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 842)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9017E213.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017E213
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.

FEATURES Location/Qualifiers
source 1. .842
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9017E213"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"
ORIGIN

Query Match 36.1%; Score 791; DB 7; Length 842;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 841; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1029 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 60

Qy       1089 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 1148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 120

Qy       1149 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 1208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGG 180

Qy       1209 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 1268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGT 240

Qy       1269 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 1328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAG 300

Qy       1329 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 360

Qy       1389 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTC 1448
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        361 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCCCCGTCTTAGAGGTGGGATGCAGATCTTC 420

Qy       1449 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 1508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAG 480

Qy       1509 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 540

Qy       1569 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 600

Qy       1629 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 660

Qy       1689 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1748
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 720
```

```
Qy       1749 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 1808
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAA 780

Qy       1809 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 1868
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 840

Qy       1869 CT 1870
             ||
Db         841 CT 842


RESULT 61
CR977355
LOCUS       CR977355                 891 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR977355 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017L184 5',
            mRNA sequence.
ACCESSION   CR977355
VERSION     CR977355.1  GI:68215676
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 891)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017L184.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017L184
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .891
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017L184"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             36.1%;  Score 791;  DB 7;  Length 891;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  791;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        513 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 60

Qy        573 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATT 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATT 120

Qy        633 CCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 180

Qy        693 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGG 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGG 240

Qy        753 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGT 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGT 300

Qy        813 GACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAG 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAG 360

Qy        873 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 420

Qy        933 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 480

Qy        993 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAG 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAG 540

Qy       1053 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 600

Qy       1113 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAG 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAG 660

Qy       1173 TCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 720

Qy       1233 ACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCA 780

Qy       1293 AAGATCCAAGA 1303
              ||||||||||| .
Db        781 AAGATCCAAGA 791


RESULT 62
CR983573
LOCUS       CR983573                 882 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR983573 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016H1014 5',
            mRNA sequence.
ACCESSION   CR983573
VERSION     CR983573.1  GI:68221907
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
```

```
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016H1014.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016H1014
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016H1014"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           35.9%;  Score 788;  DB 7;  Length 882;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  788;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 60

Qy        744 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTC 120

Qy        804 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCT 180

Qy        864 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCT 240

Qy        924 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATG 300

Qy        984 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 360

Qy       1044 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 420

Qy       1104 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 480

Qy       1164 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 540

Qy       1224 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 600

Qy       1284 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 660

Qy       1344 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 720

Qy       1404 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 780

Qy       1464 GGTAAGAC 1471
              ||||||||
Db        781 GGTAAGAC 788


RESULT 63
CT001833
LOCUS       CT001833                915 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT001833 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017H1240 5',
            mRNA sequence.
ACCESSION   CT001833
VERSION     CT001833.1  GI:68295716
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017H1240.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017H1240
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017H1240"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
```

/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 35.9%; Score 788; DB 7; Length 915;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         12 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 60

Qy         72 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 120

Qy        132 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 180

Qy        192 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 240

Qy        252 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 300

Qy        312 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 360

Qy        372 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 420

Qy        432 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 480

Qy        492 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 540

Qy        552 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 600

Qy        612 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 660

Qy        672 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 720

Qy        732 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 780

Qy        792 ACCCTTGA 799
              ||||||||
Db        781 ACCCTTGA 788
```

RESULT 64

CX872472
LOCUS CX872472 794 bp mRNA linear EST 03-FEB-2005
DEFINITION HESC4_69_D12.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone IMAGE:7487281 5', mRNA sequence.
ACCESSION CX872472
VERSION CX872472.1 GI:58555646
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 794)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM15816 row: g column: 23
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 794.
FEATURES Location/Qualifiers
source 1. .794
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7487281"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 35.9%; Score 786; DB 8; Length 794;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 68

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 128

Qy        737 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 796
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 188

Qy        797 TGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGG 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 TGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGG 248

Qy        857 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCAC 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCAC 308

Qy        917 CCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGG 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 CCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGG 368

Qy        977 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCC 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCC 428

Qy       1037 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 488

Qy       1097 CCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTA 1156
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 CCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTA 548

Qy       1157 CAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGAT 1216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 CAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGAT 608

Qy       1217 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCAT 1276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCAT 668

Qy       1277 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTT 1336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTT 728

Qy       1337 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 788

Qy       1397 AGAGTC 1402
              ||||||
Db        789 AGAGTC 794
```

```
RESULT 65
CR987122
LOCUS       CR987122               842 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR987122 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017I0338 5',
            mRNA sequence.
ACCESSION   CR987122
VERSION     CR987122.1  GI:68281007
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 842)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017I0338.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017I0338
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .842
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017I0338"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             35.9%;  Score 786;  DB 7;  Length 842;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  836;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        357 GACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 60

Qy        417 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 120

Qy        477 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 180

Qy        537 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAG 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAG 240

Qy        597 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATC 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATC 300

Qy        657 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAG 360

Qy        717 TCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 420

Qy        777 ACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCA 480

Qy        837 AAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 540

Qy        897 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 600

Qy        957 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 660

Qy       1017 ATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 1076
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGAC 720

Qy       1077 AAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 1136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 780

Qy       1137 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTC 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTC 837


RESULT 66
AL535583
LOCUS       AL535583                1037 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535583 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF011YG14 5-PRIME, mRNA sequence.
ACCESSION   AL535583
VERSION     AL535583.3  GI:45711452
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1037)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30542509.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF011BD07QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1037
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF011YG14"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             35.9%;  Score 786;  DB 1;  Length 1037;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  836;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 78

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 138

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 198

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 258

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 318

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        319 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 378

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 438

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 498

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 558

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 618

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 678

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 738

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 798

Qy        812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGA 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGA 855


RESULT 67
CR987992
LOCUS       CR987992                836 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR987992 RZPD no.9016 Homo sapiens cDNA clone RZPDp901601313 5',
            mRNA sequence.
ACCESSION   CR987992
VERSION     CR987992.1  GI:68281877
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
```

Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016O1313.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016O1313
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.

FEATURES Location/Qualifiers
source 1. .836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016O1313"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from
human T-Lymphocytes with a NotI - oligo(dT) primer [5'
GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to SalI adaptors,
digested with NotI and cloned into the NotI and SalI sites
of the pQE80LSN_cloned vector"

ORIGIN

Query Match 35.8%; Score 785; DB 7; Length 836;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 835; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        123 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 60

Qy        183 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGAC 120

Qy        243 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAA 180

Qy        303 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACC 362
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        181 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACC 240

Qy        363 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 300

Qy        423 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 360

Qy        483 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 420

Qy        543 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTC 480

Qy        603 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCC 540

Qy        663 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 600

Qy        723 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 660

Qy        783 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATC 720

Qy        843 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 780

Qy        903 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 958
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 836
```

RESULT 68
CR995740
LOCUS CR995740 872 bp mRNA linear EST 28-JUN-2005
DEFINITION CR995740 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016J076 5', mRNA sequence.
ACCESSION CR995740
VERSION CR995740.1 GI:68289625
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 872)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016J076.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016J076
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.

FEATURES Location/Qualifiers
source 1. .872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016J076"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from

human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 35.8%; Score 785; DB 7; Length 872;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy      75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy     135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy     195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy     255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy     315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy     375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy     435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy     495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy     555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy     615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy     675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy     735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy     795 CTTGA 799
           |||||
Db     781 CTTGA 785
```

RESULT 69
CR987948
LOCUS CR987948 903 bp mRNA linear EST 28-JUN-2005
DEFINITION CR987948 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017D1831 5', mRNA sequence.
ACCESSION CR987948
VERSION CR987948.1 GI:68281833
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9017D1831.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017D1831
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9017D1831"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 35.8%; Score 784; DB 7; Length 903;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      12 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 71
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATG 60

Qy      72 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 131
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGAC 120

Qy     132 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 191
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAG 180

Qy     192 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 251
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 240

Qy     252 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 311
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 CAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 300

Qy        312 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAAC 360

Qy        372 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 420

Qy        432 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 480

Qy        492 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 540

Qy        552 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAG 600

Qy        612 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAG 660

Qy        672 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 720

Qy        732 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 780

Qy        792 ACCC 795
              ||||
Db        781 ACCC 784
```

```
RESULT 70
CR991185
LOCUS       CR991185                905 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR991185 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016E0319 5',
            mRNA sequence.
ACCESSION   CR991185
VERSION     CR991185.1  GI:68285070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016E0319.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016E0319
            contact RZPD (product- support@rzpd.de) for further information.
```

```
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES            Location/Qualifiers
     source         1. .905
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="RZPDp9016E0319"
                    /tissue_type="T-Lymphocytes"
                    /dev_stage="adult"
                    /lab_host="DH10B"
                    /clone_lib="RZPD no.9016"
                    /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                    NotI; vector:
                    http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                    ; 1st strand cDNA was prepared from mRNA obtained from
                    human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                    GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. .
                    Double-stranded cDNA was ligated to SalI adaptors,
                    digested with NotI and cloned into the NotI and SalI sites
                    of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             35.8%;  Score 784;  DB 7;  Length 905;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  834;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720
```

```
Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGANCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACA 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACA 835


RESULT 71
AL535866
LOCUS       AL535866                 973 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL535866 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF013YN23 5-PRIME, mRNA sequence.
ACCESSION   AL535866
VERSION     AL535866.3  GI:45711716
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 973)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30542812.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF013CG12QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .973
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF013YN23"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             35.8%;  Score 784;  DB 1;  Length 973;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  954;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 77

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 137

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 197

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 257

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 317

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAG-G 390
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |
Db        318 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGTG 377

Qy        391 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 437

Qy        451 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 497

Qy        511 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 557

Qy        571 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 617

Qy        631 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 677

Qy        691 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 737

Qy        751 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCA 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCA 797

Qy        811 GTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACC 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 GTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACC 857

Qy        871 AGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACA 930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 AGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACA 917

Qy        931 ACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 986
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 ACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 973


RESULT 72
BM908271
LOCUS       BM908271                 1052 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6609577 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5430827
            5', mRNA sequence.
ACCESSION   BM908271
VERSION     BM908271.1  GI:19358650
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1052)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
```

```
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1902  row: b  column: 12
           High quality sequence stop: 701.
FEATURES             Location/Qualifiers
     source          1. .1052
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5430827"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           35.8%;  Score 784;  DB 3;  Length 1052;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  834;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         18 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 60

Qy         78 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 120

Qy        138 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 180

Qy        198 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 240

Qy        258 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 300

Qy        318 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 360

Qy        378 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 420

Qy        438 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 480

Qy        498 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 540

Qy        558 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 600

Qy        618 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 660

Qy        678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 720

Qy        738 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        721 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCCAGACCATCACCCTT 780

Qy        798 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 852
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 835


RESULT 73
BU153181
LOCUS       BU153181                906 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_7939215 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6008468
            5', mRNA sequence.
ACCESSION   BU153181
VERSION     BU153181.1  GI:22666713
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13193  row: f  column: 21
           High quality sequence stop: 659.
FEATURES             Location/Qualifiers
     source          1...906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6008468"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match           35.7%;  Score 783;  DB 5;  Length 906;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  783;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 60

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 120

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 180
```

```
Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 240

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 300

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 360

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 420

Qy       1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 480

Qy       1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 540

Qy       1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 600

Qy       1308 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 660

Qy       1368 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTT 720

Qy       1428 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 1487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTG 780

Qy       1488 GAG 1490
              |||
Db        781 GAG 783
```

```
RESULT 74
BQ225018
LOCUS       BQ225018                975 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7260772 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785384
            5', mRNA sequence.
ACCESSION   BQ225018
VERSION     BQ225018.1  GI:20406418
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12872  row: g  column: 17
            High quality sequence stop: 699.
```

```
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5785384"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN
```

```
  Query Match             35.7%;  Score 783;  DB 3;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  783;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 60

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 120

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 180

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 240

Qy        257 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 300

Qy        317 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 360

Qy        377 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 420

Qy        437 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 480

Qy        497 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 540

Qy        557 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 600

Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 660

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 720

Qy        737 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 780

Qy        797 TGA 799
              |||
Db        781 TGA 783
```

```
RESULT 75
CD244277
LOCUS       CD244277                825 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14096361 NIH_MGC_180 Homo sapiens cDNA clone
            IMAGE:30377981 5', mRNA sequence.
ACCESSION   CD244277
VERSION     CD244277.1  GI:31004741
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 825)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM443  row: b  column: 06
            High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .825
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30377981"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_180"
                     /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI;
                     Site_2: EcoRV (destroyed); Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.68 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match           35.7%;  Score 782;  DB 6;  Length 825;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy        796 TT 797
              ||
Db        781 TT 782


RESULT 76
BQ421336
LOCUS       BQ421336                859 bp    mRNA    linear   EST 23-MAY-2002
DEFINITION  AGENCOURT_7772249 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014685
            5', mRNA sequence.
ACCESSION   BQ421336
VERSION     BQ421336.1  GI:21116651
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13209  row: i  column: 22
            High quality sequence stop: 670.
FEATURES             Location/Qualifiers
     source          1. .859
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6014685"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
```

Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 35.7%; Score 782; DB 5; Length 859;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy          76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy         136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy         196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy         376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy         436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy         496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy         556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy         616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy         676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy         736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780

Qy         796 TT 797
               ||
Db         781 TT 782
```

RESULT 77
CX163961
LOCUS CX163961 781 bp mRNA linear EST 23-DEC-2004
DEFINITION HESC2_11_B03.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone IMAGE:7467223 5', mRNA sequence.
ACCESSION CX163961
VERSION CX163961.1 GI:56794041
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM15764 row: d column: 05
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 781.
FEATURES Location/Qualifiers
source 1. .781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7467223"
/sex="male"
/tissue_type="human embryonic stem cells differentiated to an early endodermal cell type"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_258"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early endodermal cell type. Cell line id and NIH Registry designation is BG01. Positive for GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP expression. Passage number 40. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."
ORIGIN

Query Match 35.6%; Score 781; DB 8; Length 781;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 781; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        1193 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 1252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 60

Qy        1253 CGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 120

Qy        1313 CATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCAC 1372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCAC 180
```

```
Qy       1373 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGG 1432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGG 240

Qy       1433 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCC 1492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCC 300

Qy       1493 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1552
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 360

Qy       1553 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 420

Qy       1613 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 480

Qy       1673 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 1732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCAT 540

Qy       1733 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 1792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTT 600

Qy       1793 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 1852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAA 660

Qy       1853 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 1912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 720

Qy       1913 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 1972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 780

Qy       1973 G 1973
             |
Db        781 G 781
```

RESULT 78
CT001360
LOCUS CT001360 799 bp mRNA linear EST 28-JUN-2005
DEFINITION CT001360 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016K1121 5', mRNA sequence.
ACCESSION CT001360
VERSION CT001360.1 GI:68295243
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 799)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016K1121.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016K1121
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .799
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016K1121"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"
ORIGIN

Query Match 35.6%; Score 781; DB 7; Length 799;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 781; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 C 795
              |
Db        781 C 781


RESULT 79
CR988748
LOCUS       CR988748                880 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR988748 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017P1023 5',
            mRNA sequence.
ACCESSION   CR988748
VERSION     CR988748.1  GI:68282633
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 880)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017P1023.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017P1023
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .880
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9017P1023"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           35.6%;  Score 781;  DB 7;  Length 880;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  781;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        528 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGAC 60

Qy        588 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAG 120

Qy        648 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 180

Qy        708 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 240

Qy        768 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAAC 300

Qy        828 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTT 360

Qy        888 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCT 420

Qy        948 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 480

Qy       1008 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 540

Qy       1068 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAG 600

Qy       1128 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTG 660

Qy       1188 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 720

Qy       1248 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 780

Qy       1308 G 1308
              |
Db        781 G 781


RESULT 80
BU155752
LOCUS       BU155752                857 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_7977748 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6164325
            5', mRNA sequence.
ACCESSION   BU155752
VERSION     BU155752.1  GI:22669284
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 857)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13521  row: d  column: 22
            High quality sequence stop: 722.
FEATURES             Location/Qualifiers
     source          1. .857
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6164325"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match           35.6%;  Score 780;  DB 5;  Length 857;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  780;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 780


RESULT 81
BU180938
LOCUS       BU180938                875 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7943922 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6009261
            5', mRNA sequence.
ACCESSION   BU180938
VERSION     BU180938.1  GI:22694922
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 875)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/:
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13195  row: g  column: 22
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .875
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6009261"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match           35.6%;  Score 780;  DB 5;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  780;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 60

Qy         87 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 120

Qy        147 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 180
```

```
Qy        207 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 240

Qy        267 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 300

Qy        327 AAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATC 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATC 360

Qy        387 CAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 420

Qy        447 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 480

Qy        507 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 540

Qy        567 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 600

Qy        627 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 660

Qy        687 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 720

Qy        747 GGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAG 806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAG 780


RESULT 82
BQ880781
LOCUS       BQ880781                924 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8229020 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6183750 5', mRNA sequence.
ACCESSION   BQ880781
VERSION     BQ880781.1  GI:22272789
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 924)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13571  row: n  column: 07
            High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .924
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6183750"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           35.6%;  Score 780;  DB 5;  Length 924;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  780;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 60

Qy         78 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 120

Qy        138 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 180

Qy        198 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 240

Qy        258 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 300

Qy        318 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 360

Qy        378 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 420

Qy        438 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 480

Qy        498 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 540

Qy        558 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 600

Qy        618 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 660

Qy        678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 720

Qy        738 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        721 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 780


RESULT 83
BU149754
LOCUS       BU149754                854 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8124349 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6178099 5', mRNA sequence.
ACCESSION   BU149754
VERSION     BU149754.1  GI:22663286
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 854)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13557  row: b  column: 20
            High quality sequence stop: 705.
FEATURES             Location/Qualifiers
     source          1. .854
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6178099"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           35.5%;  Score 779;  DB 5;  Length 854;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  779;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 60

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 120

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 180

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 240

Qy        257 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 300

Qy        317 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 360

Qy        377 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 420

Qy        437 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 480

Qy        497 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 540

Qy        557 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 600

Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 660

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 720

Qy        737 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 779


RESULT 84
AL560464
LOCUS       AL560464                897 bp    mRNA    linear   EST 02-APR-2004
DEFINITION  AL560464 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
            clone CS0DG006YF03 5-PRIME, mRNA sequence.
ACCESSION   AL560464
VERSION     AL560464.3  GI:46185835
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:31284595.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DG006CC02QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

/clone="CS0DG006YF03"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 35.5%; Score 779; DB 1; Length 897;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 879; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCYGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTTCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 895
              |||||||||||||||||||||||||||||||||||||||||
```

```
Db        841 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 881
```

RESULT 85
BU172945
LOCUS BU172945 899 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7964816 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6169483 5', mRNA sequence.
ACCESSION BU172945
VERSION BU172945.1 GI:22686929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 899)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13534 row: k column: 20
High quality sequence stop: 665.

FEATURES Location/Qualifiers
source 1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6169483"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 35.5%; Score 779; DB 5; Length 899;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 779; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 60

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 120

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 180

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 240

Qy        257 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 300

Qy        317 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 360
```

```
Qy        377 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 420

Qy        437 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 480

Qy        497 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 540

Qy        557 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 600

Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 660

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 720

Qy        737 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 779
```

RESULT 86
CR979926
LOCUS CR979926 906 bp mRNA linear EST 24-JUN-2005
DEFINITION CR979926 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016H072 5', mRNA sequence.
ACCESSION CR979926
VERSION CR979926.1 GI:68218260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016H072.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016H072
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016H072"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"
ORIGIN

Query Match 35.4%; Score 776; DB 7; Length 906;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 826; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        681 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 60

Qy        741 CTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAG 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAG 120

Qy        801 GTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATT 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATT 180

Qy        861 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTG 240

Qy        921 TCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 300

Qy        981 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGT 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGT 360

Qy       1041 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 420

Qy       1101 CAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAAC 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAAC 480

Qy       1161 ATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 540

Qy       1221 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAG 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAG 600

Qy       1281 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATC 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATC 660

Qy       1341 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 720

Qy       1401 TCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 1460
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        721 TCCACCCTGCACCTGGTGCTCCGTCTTANAGGTGGGATGCAGATCTTCGTGAAGACCCTG 780

Qy       1461 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 827
```

```
RESULT 87
BQ927751
LOCUS       BQ927751                919 bp    mRNA    linear   EST 20-AUG-2002
DEFINITION  AGENCOURT_8822246 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381542
            5', mRNA sequence.
ACCESSION   BQ927751
VERSION     BQ927751.1  GI:22342782
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 919)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2571  row: o  column: 15
            High quality sequence stop: 732.
FEATURES             Location/Qualifiers
     source          1. .919
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6381542"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_40"
                     /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California,  Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life  Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             35.4%;  Score 775;  DB 5;  Length 919;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  825;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         18 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 60

Qy         78 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 120

Qy        138 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 180

Qy        198 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 240

Qy        258 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 300

Qy        318 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 360

Qy        378 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 420

Qy        438 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 480

Qy        498 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 540

Qy        558 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 600

Qy        618 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 660

Qy        678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 720

Qy        738 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        721 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAACATCACCCTT 780

Qy        798 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
              |||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 826


RESULT 88
BQ643348
LOCUS       BQ643348                920 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8485464 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:6305249
            5', mRNA sequence.
ACCESSION   BQ643348
VERSION     BQ643348.1  GI:21767520
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 920)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2527  row: d  column: 18
            High quality sequence stop: 679.
FEATURES             Location/Qualifiers
     source          1. .920
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6305249"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
```

/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 35.3%; Score 774; DB 5; Length 920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1067 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACA 1126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACA 60

Qy       1127 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 1186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 120

Qy       1187 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 180

Qy       1247 CACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 240

Qy       1307 GGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 300

Qy       1367 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 360

Qy       1427 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 1486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 420

Qy       1487 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 480

Qy       1547 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 540

Qy       1607 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 1666
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 600

Qy       1667 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 1726
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 660

Qy       1727 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 1786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCA 720

Qy       1787 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 774
```

RESULT 89
BM463129
LOCUS BM463129 1070 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6447311 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5537833 5', mRNA sequence.
ACCESSION BM463129
VERSION BM463129.1 GI:18512169
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1070)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM12229 row: m column: 02
High quality sequence stop: 667.
FEATURES Location/Qualifiers
source 1. .1070
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5537833"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
ORIGIN

Query Match 35.3%; Score 774; DB 3; Length 1070;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 60

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 120

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 180

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 240

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 300

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 360

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 420

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 480
```

```
Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 540

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 600

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 660

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 720

Qy        742 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 774


RESULT 90
CX871255
LOCUS       CX871255                773 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_55_A10.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7485621 5', mRNA sequence.
ACCESSION   CX871255
VERSION     CX871255.1  GI:58554429
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15812  row: b  column: 19
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 773.
FEATURES             Location/Qualifiers
     source          1. .773
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7485621"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match           35.3%;  Score 773;  DB 8;  Length 773;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  773;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        956 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 60

Qy       1016 CATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 120

Qy       1076 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 180

Qy       1136 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCG 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCG 240

Qy       1196 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 300

Qy       1256 GGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 360

Qy       1316 CCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 420

Qy       1376 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 480

Qy       1436 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAG 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAG 540

Qy       1496 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 600

Qy       1556 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 660

Qy       1616 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 720

Qy       1676 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACA 773


RESULT 91
BU902337
LOCUS       BU902337                1098 bp    mRNA    linear   EST 17-OCT-2002
DEFINITION  AGENCOURT_10127301 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6503328
            5', mRNA sequence.
ACCESSION   BU902337
```

```
VERSION     BU902337.1  GI:24084250
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1098)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14060  row: n  column: 01
            High quality sequence stop: 810.
FEATURES             Location/Qualifiers
     source          1. .1098
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6503328"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match            35.3%;  Score 773;  DB 5;  Length 1098;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches  873;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 65

Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 125

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         126 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 185

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         186 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 245

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         246 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 305

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         306 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 365

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         366 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 425

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         426 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 485

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         486 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 545

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         546 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 605

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         606 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 665

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         666 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 725

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         726 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 785

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         786 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 845

Qy       1863 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 1897
             ||||||||||||||||||||||||||||||||||||
Db         846 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 880


RESULT 92
BQ431211
LOCUS       BQ431211                905 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AGENCOURT_7751282 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062447
            5', mRNA sequence.
ACCESSION   BQ431211
VERSION     BQ431211.1  GI:21170287
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13333  row: o  column: 24
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1. .905
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6062447"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match            35.2%;  Score 771;  DB 5;  Length 905;
  Best Local Similarity  100.0%;  Pred. No. 0;
```

```
  Matches  771;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1165 AGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 1224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 60

Qy       1225 AGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATG 1284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATG 120

Qy       1285 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTG 1344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTG 180

Qy       1345 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 240

Qy       1405 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 300

Qy       1465 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 360

Qy       1525 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 420

Qy       1585 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 480

Qy       1645 TGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 540

Qy       1705 CTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1764
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 600

Qy       1765 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1824
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 660

Qy       1825 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 1884
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 720

Qy       1885 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 771


RESULT 93
CR978349
LOCUS       CR978349                926 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR978349 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016N154 5',
            mRNA sequence.
ACCESSION   CR978349
VERSION     CR978349.1  GI:68216683
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 926)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016N154.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016N154
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .926
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016N154"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             35.2%;  Score 771;  DB 7;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  771;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         14 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 73
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 63

Qy         74 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 123

Qy        134 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 183

Qy        194 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 243

Qy        254 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 303

Qy        314 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 363

Qy        374 CAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGC 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGC 423

Qy        434 CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTAC 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTAC 483
```

```
Qy        494 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 CCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 543

Qy        554 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 TAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGAT 603

Qy        614 CCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCT 673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 CCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCT 663

Qy        674 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 GGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGT 723

Qy        734 ACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 ACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 774


RESULT 94
BQ051425
LOCUS       BQ051425                1038 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6763507 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:5792842
            5', mRNA sequence.
ACCESSION   BQ051425
VERSION     BQ051425.1  GI:19810765
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1038)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12891  row: n  column: 11
            High quality sequence stop: 656.
FEATURES             Location/Qualifiers
     source          1. .1038
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5792842"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match             35.2%;  Score 771;  DB 3;  Length 1038;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  771;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 60

Qy         89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 120

Qy        149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 180

Qy        209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 240

Qy        269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 300

Qy        329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 360

Qy        389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 420

Qy        449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 480

Qy        509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 540

Qy        569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 600

Qy        629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 660

Qy        689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 720

Qy        749 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 771


RESULT 95
CX873532
LOCUS       CX873532                770 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_82_A09.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7488284 5', mRNA sequence.
ACCESSION   CX873532
VERSION     CX873532.1  GI:58556706
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 770)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
```

```
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
           University of Georgia
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM15819  row: a  column: 18
           Seq primer: JENREV (CAGGAAACAGCTATGACC)
           High quality sequence stop: 770.
FEATURES             Location/Qualifiers
     source          1. .770
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7488284"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             35.1%;  Score 770;  DB 8;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  770;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        591 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGG 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGG 60

Qy        651 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 120

Qy        711 AAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 180

Qy        771 ACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTC 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTC 240

Qy        831 AAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 300

Qy        891 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 360

Qy        951 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 420

Qy       1011 AAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 480

Qy       1071 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 540

Qy       1131 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTG 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTG 600

Qy       1191 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 660

Qy       1251 CTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 720

Qy       1311 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGA 770


RESULT 96
BU175290
LOCUS       BU175290                897 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7939019 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6008360
            5', mRNA sequence.
ACCESSION   BU175290
VERSION     BU175290.1  GI:22689274
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13193  row: b  column: 09
           High quality sequence stop: 597.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6008360"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match             35.1%;  Score 769;  DB 5;  Length 897;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  769;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 60

Qy         89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
```

```
Db        61 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 120

Qy       149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208

Db       121 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 180

Qy       209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268

Db       181 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 240

Qy       269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328

Db       241 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 300

Qy       329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388

Db       301 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 360

Qy       389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448

Db       361 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 420

Qy       449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508

Db       421 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 480

Qy       509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568

Db       481 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 540

Qy       569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628

Db       541 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 600

Qy       629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688

Db       601 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 660

Qy       689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 748

Db       661 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 720

Qy       749 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797

Db       721 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 769


RESULT 97
BU521200
LOCUS       BU521200                878 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT_10167694 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6527033
            5', mRNA sequence.
ACCESSION   BU521200
VERSION     BU521200.1  GI:22828726
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14123  row: e  column: 17
            High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6527033"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match             35.0%;  Score 767;  DB 5;  Length 878;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  767;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75

Db         1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy        76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135

Db        61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy       136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195

Db       121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy       196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255

Db       181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy       256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315

Db       241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy       316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375

Db       301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy       376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435

Db       361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy       436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495

Db       421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy       496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555

Db       481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy       556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615

Db       541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy       616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675

Db       601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy       676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735

Db       661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy       736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 782

Db       721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 767
```

```
RESULT 98
BQ879588
LOCUS       BQ879588                 917 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8232277 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6182678 5', mRNA sequence.
ACCESSION   BQ879588
VERSION     BQ879588.1  GI:22271596
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13569  row: a  column: 15
            High quality sequence stop: 654.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6182678"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match             35.0%;  Score 767;  DB 5;  Length 917;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  767;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 60

Qy         91 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 120

Qy        151 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 180

Qy        211 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 240

Qy        271 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 300

Qy        331 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 360

Qy        391 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 420

Qy        451 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 480

Qy        511 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 540

Qy        571 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 600

Qy        631 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 660

Qy        691 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 720

Qy        751 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 767


RESULT 99
BQ723637
LOCUS       BQ723637                 977 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8229565 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184595 5', mRNA sequence.
ACCESSION   BQ723637
VERSION     BQ723637.1  GI:21862534
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 977)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13574  row: a  column: 12
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .977
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6184595"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
```

/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 35.0%; Score 767; DB 5; Length 977;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 85

Qy      75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 145

Qy     135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 205

Qy     195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     206 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 265

Qy     255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     266 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 325

Qy     315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     326 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 385

Qy     375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     386 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 445

Qy     435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     446 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 505

Qy     495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     506 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 565

Qy     555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     566 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 625

Qy     615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     626 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 685

Qy     675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     686 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 745

Qy     735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 781
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     746 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGG 792
```

RESULT 100
BM803902

```
LOCUS       BM803902                1042 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  AGENCOURT_6439424 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5520601
            5', mRNA sequence.
ACCESSION   BM803902
VERSION     BM803902.1  GI:19120725
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1042)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12184  row: o  column: 02
            High quality sequence stop: 687.
FEATURES             Location/Qualifiers
     source          1. .1042
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5520601"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN
```

Query Match 35.0%; Score 767; DB 3; Length 1042;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 60

Qy      89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 120

Qy     149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 180

Qy     209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 240

Qy     269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 300

Qy     329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 360

Qy     389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 420

Qy     449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508
```

```
Db        421 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 480

Qy        509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568

Db        481 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 540

Qy        569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628

Db        541 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 600

Qy        629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688

Db        601 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 660

Qy        689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 748

Db        661 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 720

Qy        749 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795

Db        721 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 767
```

```
RESULT 101
BM464473
LOCUS       BM464473                1080 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6438785 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535560
            5', mRNA sequence.
ACCESSION   BM464473
VERSION     BM464473.1  GI:18513515
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12223  row: n  column: 09
            High quality sequence stop: 702.
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5535560"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match             35.0%;  Score 767;  DB 3;  Length 1080;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  767;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         65 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 124

Db         20 GACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCC 79
```

```
Qy        125 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 184

Db         80 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGA 139

Qy        185 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 244

Db        140 CCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTA 199

Qy        245 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 304

Db        200 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAAT 259

Qy        305 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 364

Db        260 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCAT 319

Qy        365 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 424

Db        320 CGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 379

Qy        425 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 484

Db        380 GATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 439

Qy        485 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 544

Db        440 AGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGAC 499

Qy        545 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 604

Db        500 CCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAA 559

Qy        605 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 664

Db        560 GGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGG 619

Qy        665 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724

Db        620 AAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 679

Qy        725 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 784

Db        680 GCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 739

Qy        785 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCA 831

Db        740 GACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCA 786
```

```
RESULT 102
BQ878048
LOCUS       BQ878048                908 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8137209 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184023 5', mRNA sequence.
ACCESSION   BQ878048
VERSION     BQ878048.1  GI:22270056
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
```

```
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13572  row: i  column: 16
            High quality sequence stop: 673.
FEATURES             Location/Qualifiers
     source          1. .908
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6184023"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           34.9%;  Score 766;  DB 5;  Length 908;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  816;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 75

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 135

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 195

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 255

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 315

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        316 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 375

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 435

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 495

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 555

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 615

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 675

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 735

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 795

Qy        812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              |||||||||||||||||||||||||||||||||||||
Db        796 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 832


RESULT 103
BQ880336
LOCUS       BQ880336                924 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8122459 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6179541 5', mRNA sequence.
ACCESSION   BQ880336
VERSION     BQ880336.1  GI:22272344
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 924)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13560  row: n  column: 22
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .924
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6179541"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           34.9%;  Score 766;  DB 5;  Length 924;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  816;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         22 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 81

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 141

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 201

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 261

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 321

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        322 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 381

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 441

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 501

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 561

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 621

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 681

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 741

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 801

Qy        812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              |||||||||||||||||||||||||||||||||||||
Db        802 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 838


RESULT 104
CR996383
LOCUS       CR996383                 897 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR996383 RZPD no.9017 Homo sapiens cDNA clone RZPDp90170162 5',
            mRNA sequence.
ACCESSION   CR996383
VERSION     CR996383.1  GI:68290268
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp90170162.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp90170162
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .897
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp90170162"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. .
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.9%;  Score 765;  DB 7;  Length 897;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  865;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 75

Qy       1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 135

Qy       1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 195

Qy       1506 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 255

Qy       1566 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 315

Qy       1626 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 375

Qy       1686 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1745
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        376 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTCAAG 435

Qy       1746 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 495
```

```
Qy       1806 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1865
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 555

Qy       1866 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1925
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        556 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACCCTGACTGGTAAG 615

Qy       1926 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 1985
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 675

Qy       1986 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 2045
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAA 735

Qy       2046 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 2105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTG 795

Qy       2106 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 2165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        796 CGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTT 855

Qy       2166 TCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||
Db        856 TCCTTTCAATAAAGTTGTTGCATTCCC 882
```

```
RESULT 105
BQ894538
LOCUS       BQ894538                900 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8119193 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6179593 5', mRNA sequence.
ACCESSION   BQ894538
VERSION     BQ894538.1  GI:22286552
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13561  row: a  column: 02
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6179593"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN
```

```
  Query Match           34.9%;  Score 765;  DB 5;  Length 900;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  765;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 60

Qy         91 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 120

Qy        151 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 180

Qy        211 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 240

Qy        271 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 300

Qy        331 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 360

Qy        391 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 420

Qy        451 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 480

Qy        511 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 540

Qy        571 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 600

Qy        631 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 660

Qy        691 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 720

Qy        751 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 795
              |||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCC 765
```

```
RESULT 106
CF552130
LOCUS       CF552130                892 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15596110 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30530584 5', mRNA sequence.
ACCESSION   CF552130
VERSION     CF552130.1  GI:34888964
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM619  row: p  column: 17
            High quality sequence stop: 699.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30530584"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match             34.8%;  Score 763;  DB 6;  Length 892;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  813;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        464 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 109

Qy        524 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAG 169

Qy        584 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 TGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCA 229

Qy        644 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 289

Qy        704 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 CATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTT 349

Qy        764 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 CGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 409

Qy        824 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 GAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGAT 469

Qy        884 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 CTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 529

Qy        944 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 GTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 589

Qy       1004 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 GACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 649

Qy       1064 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        650 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 709

Qy       1124 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 769

Qy       1184 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1243
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        770 CCTGGTGCTCCGTCTCAAAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 829

Qy       1244 CATCACTCTCGAGGTGGAGCCGAGTGACACCATT 1277
              ||||||||||||||||||||||||||||||||||
Db        830 CATCACTCTCGAGGTGGAGCCGAGTGACACCATT 863


RESULT 107
BQ226311
LOCUS       BQ226311                911 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7511081 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6055218
            5', mRNA sequence.
ACCESSION   BQ226311
VERSION     BQ226311.1  GI:20407711
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13315  row: b  column: 19
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6055218"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             34.8%;  Score 763;  DB 3;  Length 911;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  763;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 70

Qy        84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        71 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 130

Qy       144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       131 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 190

Qy       204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       191 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 250

Qy       264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       251 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 310

Qy       324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       311 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 370

Qy       384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       371 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 430

Qy       444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       431 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 490

Qy       504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       491 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 550

Qy       564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       551 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 610

Qy       624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       611 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 670

Qy       684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       671 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 730

Qy       744 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 786
             |||||||||||||||||||||||||||||||||||||||||||
Db       731 AGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 773
```

RESULT 108
BM458528

```
LOCUS       BM458528               1133 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6414225 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5497173
            5', mRNA sequence.
ACCESSION   BM458528
VERSION     BM458528.1  GI:18507568
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1133)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12126  row: n  column: 22
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .1133
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5497173"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 34.8%; Score 763; DB 3; Length 1133;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       475 ACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        18 ACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 77

Qy       535 TCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCG 594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        78 TCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCG 137

Qy       595 AGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGA 654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       138 AGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGA 197

Qy       655 TCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAG 714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       198 TCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAG 257

Qy       715 AGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACAC 774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       258 AGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACAC 317

Qy       775 TCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAG 834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       318 TCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAG 377

Qy       835 CAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       378 CAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 437

Qy       895 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       438 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 497

Qy       955 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       498 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 557

Qy      1015 CCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 1074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       558 CCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAG 617

Qy      1075 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 1134
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       618 ACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 677

Qy      1135 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCC 1194
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCC 737

Qy       1195 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1237
                ||||||||||||||||||||||||||||||||||||||||||||
Db        738 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 780


RESULT 109
DR763328
LOCUS       DR763328                761 bp    mRNA    linear   EST 21-JUL-2005
DEFINITION  HESC4_151_C07.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7973103 5', mRNA sequence.
ACCESSION   DR763328
VERSION     DR763328.1  GI:71056028
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 761)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM 17092  row: f  column: 13
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 761.
FEATURES             Location/Qualifiers
     source          1. .761
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7973103"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             34.7%;  Score 761;  DB 8;  Length 761;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  761;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        946 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 60

Qy       1006 CTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAA 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAA 120

Qy       1066 AGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAAC 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAAC 180

Qy       1126 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 240

Qy       1186 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 1245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 300

Qy       1246 TCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 1305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 360

Qy       1306 AGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 1365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATG 420

Qy       1366 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTC 480

Qy       1426 TTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAG 1485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAG 540

Qy       1486 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 600

Qy       1546 CTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 660

Qy       1606 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 1665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 720

Qy       1666 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1706
              |||||||||||||||||||||||||||||||||||||||||
Db        721 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 761


RESULT 110
CD518820
LOCUS       CD518820                889 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14375891 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30408830 5', mRNA sequence.
ACCESSION   CD518820
VERSION     CD518820.1  GI:31450538
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 889)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
```

```
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM492  row: g  column: 15
            High quality sequence stop: 692.
FEATURES              Location/Qualifiers
     source           1. .889
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:30408830"
                      /tissue_type="White Matter"
                      /dev_stage="Unknown"
                      /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                      /clone_lib="NIH_MGC_181"
                      /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                      (destroyed); Library is oligo-dT primed and directionally
                      cloned (EcoRV site is destroyed upon cloning). Average
                      insert size 1.42 kb. Library was constructed by
                      (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             34.7%;  Score 761;  DB 6;  Length 889;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  761;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 60

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 120

Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 180

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 240

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 300

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 360

Qy        740 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGA 420

Qy        800 GGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCAT 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCAT 480

Qy        860 TCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCT 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCT 540

Qy        920 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 600

Qy        980 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 660

Qy       1040 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 720

Qy       1100 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              |||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 761


RESULT 111
CR997819
LOCUS       CR997819                908 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR997819 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017A094 5',
            mRNA sequence.
ACCESSION   CR997819
VERSION     CR997819.1  GI:68291704
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017A094.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017A094
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .908
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9017A094"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. .
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.7%;  Score 761;  DB 7;  Length 908;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  861;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          46 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 105

Qy         126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         106 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 165

Qy         186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         166 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 225

Qy         246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         226 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 285

Qy         306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
                |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         286 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 345

Qy         366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         346 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 405

Qy         426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         406 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 465

Qy         486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         466 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 525

Qy         546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         526 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 585

Qy         606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         586 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 645

Qy         666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         646 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 705

Qy         726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         706 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 765

Qy         786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         766 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 825

Qy         846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         826 GACAAGGAAGGCATTCCTCCTGACCAGCAGANGTTGATCTTTGCCGGAAAGCAGCTGGAA 885

Qy         906 GATGGGCGCACCCTGTCTGACTA 928
                |||||||||||||||||||||||
Db         886 GATGGGCGCACCCTGTCTGACTA 908


RESULT 112
CX871258
LOCUS       CX871258                760 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_55_B01.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7485651 5', mRNA sequence.
ACCESSION   CX871258
VERSION     CX871258.1  GI:58554432
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15812  row: d  column: 01
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 760.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7485651"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             34.7%;  Score 760;  DB 8;  Length 760;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  760;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          10 CGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAA 69
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAA 60

Qy          70 TGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTG 129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTG 120

Qy         130 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGC 189
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGC 180

Qy         190 AGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACA 249
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACA 240

Qy         250 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCG 309
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCG 300

Qy         310 TGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGA 369
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGA 360

Qy        370 ACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCT 429
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCT 420

Qy        430 TTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 489
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 480

Qy        490 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 549
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGA 540

Qy        550 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 609
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAA 600

Qy        610 AGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAAC 669
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAAC 660

Qy        670 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 729
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 720

Qy        730 TGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 769
                    |||||||||||||||||||||||||||||||||||||
Db        721 TGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 760


RESULT 113
CT004188
LOCUS       CT004188                908 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT004188 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016H031 5',
            mRNA sequence.
ACCESSION   CT004188
VERSION     CT004188.1  GI:68298071
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016H031.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016H031
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .908
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016H031"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           34.7%;  Score 760;  DB 7;  Length 908;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  810;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 114

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 174

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 234

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 294

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        295 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 354

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 414

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 474

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 534

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 594

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 654

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 714

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 774

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 834

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGA 876
              ||||||||||||||||||||||||||||||
```

```
Db        835 GACAAGGAAGGCATTCCTCCTGACCAGCAGA 865


RESULT 114
CR999605
LOCUS       CR999605                917 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR999605 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017H015 5',
            mRNA sequence.
ACCESSION   CR999605
VERSION     CR999605.1  GI:68293490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017H015.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017H015
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017H015"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']).
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.7%;  Score 760;  DB 7;  Length 917;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  860;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1023 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 60

Qy       1083 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 1142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 120

Qy       1143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 1202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGA 180

Qy       1203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 1262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAG 240

Qy       1263 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 300

Qy       1323 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1382
             || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 360

Qy       1383 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 1442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAG 420

Qy       1443 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 1502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACC 480

Qy       1503 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 540

Qy       1563 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 1622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 600

Qy       1623 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 1682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAG 660

Qy       1683 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        661 ACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATCGAGAATGTC 720

Qy       1743 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCT 780

Qy       1803 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 840

Qy       1863 CTGCACCTGGTGCTCCGTCTCA 1884
             |||||||||||||||||||||||
Db        841 CTGCACCTGGTGCTCCGTCTCA 862


RESULT 115
BX438063
LOCUS       BX438063                1081 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX438063 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA003YM22 5-PRIME, mRNA sequence.
ACCESSION   BX438063
VERSION     BX438063.2  GI:47006118
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1081)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30777624.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
```

```
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DA003BG11QP1&c=6328.r.
FEATURES              Location/Qualifiers
     source           1. .1081
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="CS0DA003YM22"
                      /tissue_type="NEUROBLASTOMA"
                      /clone_lib="Homo sapiens NEUROBLASTOMA"
                      /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                      with a NotI-oligo(dT) primer. Five prime end enriched,
                      double-strand cDNA was digested with Not I and cloned into
                      the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                      Library was not normalized."
ORIGIN

  Query Match            34.7%;  Score 760;  DB 5;  Length 1081;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  810;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 93

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 153

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 213

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 273

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        274 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 333

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 393

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 453

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 513

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 573

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 633

Qy        649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 693

Qy        709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        694 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 753

Qy        769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 813

Qy        829 TCAAAGCAAAGATCCAAGACAAGGAAGGCAT 859
              |||||||||||||||||||||||||||||||
Db        814 TCAAAGCAAAGATCCAAGACAAGGAAGGCAT 844


RESULT 116
CF552441
LOCUS       CF552441                 862 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15608844 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30530213 5', mRNA sequence.
ACCESSION   CF552441
VERSION     CF552441.1  GI:34889275
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 862)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM619  row: a  column: 06
            High quality sequence stop: 706.
FEATURES              Location/Qualifiers
     source           1. .862
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:30530213"
                      /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                      /clone_lib="NIH_MGC_183"
                      /note="Organ: Pooled muscle (cardiac and skeletal);
                      Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                      NotI; Library is oligo-dT primed and directionally cloned
                      (EcoRV site is destroyed upon cloning). Average insert
                      size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match            34.6%;  Score 759;  DB 6;  Length 862;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  759;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        331 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 60

Qy        391 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 120

Qy        451 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 180
```

```
Qy        511 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 240

Qy        571 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 300

Qy        631 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 360

Qy        691 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTG 420

Qy        751 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCA 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCA 480

Qy        811 GTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACC 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACC 540

Qy        871 AGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACA 930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACA 600

Qy        931 ACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 660

Qy        991 TCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTG 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTG 720

Qy       1051 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1089
              |||||||||||||||||||||||||||||||||||||||
Db        721 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 759


RESULT 117
BX418049
LOCUS       BX418049               1030 bp    mRNA    linear   EST 03-MAY-2004
DEFINITION  BX418049 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF001YG07 5-PRIME, mRNA sequence.
ACCESSION   BX418049
VERSION     BX418049.2  GI:46958378
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1030)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30769454.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF001AD04QP2&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1030
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF001YG07"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             34.6%;  Score 759;  DB 5;  Length 1030;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  879;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy         14 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 73
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCA 60

Qy         74 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACAC 120

Qy        134 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAG 180

Qy        194 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 240

Qy        254 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAA 300

Qy        314 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGT 360

Qy        374 CAAAGCAAAGATCCAG-GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGCAAAGATCCAGAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 420

Qy        433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 480

Qy        493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 540

Qy        553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 600

Qy        613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 660

Qy        673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 720

Qy        733 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCA 780

Qy        793 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGG 840
```

```
Qy        853 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 892
              ||||||||||||||||||||||||||||||||||||||||
Db        841 AAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 880


RESULT 118
BM557788
LOCUS       BM557788               1057 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6588330 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5477723
            5', mRNA sequence.
ACCESSION   BM557788
VERSION     BM557788.1  GI:18800045
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1057)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1996  row: d  column: 12
            High quality sequence stop: 696.
FEATURES             Location/Qualifiers
     source          1. .1057
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5477723"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             34.6%;  Score 759;  DB 3;  Length 1057;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  809;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 60

Qy        769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 120

Qy        829 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 180

Qy        889 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 240

Qy        949 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 300

Qy       1009 GTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 360

Qy       1069 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 420

Qy       1129 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 480

Qy       1189 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 540

Qy       1249 CTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 600

Qy       1309 AAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGAC 660

Qy       1369 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTA 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTA 720

Qy       1429 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGG 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGG 780

Qy       1489 AGCCGAGTGACACCATTGAGAATGTCAAGG 1518
              |||||||||||||||||||||||||||||
Db        781 AGCCGAGTGACACCATTGAGAATGTCAAGG 810


RESULT 119
CR987157
LOCUS       CR987157               880 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR987157 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016M0810 5',
            mRNA sequence.
ACCESSION   CR987157
VERSION     CR987157.1  GI:68281042
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 880)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016M0810.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
```

```
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016M0810
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .880
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9016M0810"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9016"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.6%;  Score 758;  DB 7;  Length 880;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  808;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 102

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 162

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 222

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 282

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        283 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 342

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 402

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 462

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 522

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 582

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 642

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 702

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 762

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 822

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCA 874
              |||||||||||||||||||||||||||||
Db        823 GACAAGGAAGGCATTCCTCCTGACCAGCA 851


RESULT 120
CR976288
LOCUS       CR976288                931 bp    mRNA    linear   EST 22-JUN-2005
DEFINITION  CR976288 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016P0715 5',
            mRNA sequence.
ACCESSION   CR976288
VERSION     CR976288.1  GI:68146242
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 931)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016P0715.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016P0715
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .931
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9016P0715"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9016"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.6%;  Score 758;  DB 7;  Length 931;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  858;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
```

```
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCNCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 780

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        781 CTTGAGGTCGAGCCCAGTCACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 840

Qy        855 GGCATTCCTCCTGACCAGCA 874
               ||||||||||||||||||||
Db        841 GGCATTCCTCCTGACCAGCA 860
```

RESULT 121
BU160892

```
LOCUS       BU160892                861 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7969863 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6147171
            5', mRNA sequence.
ACCESSION   BU160892
VERSION     BU160892.1  GI:22674802
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13476  row: j  column: 04
            High quality sequence stop: 613.
FEATURES             Location/Qualifiers
     source          1..861
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6147171"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN
```

Query Match 34.5%; Score 757; DB 5; Length 861;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600
```

```
Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 772
              |||||||||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 757


RESULT 122
CT000802
LOCUS       CT000802                908 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT000802 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017E1726 5',
            mRNA sequence.
ACCESSION   CT000802
VERSION     CT000802.1  GI:68294685
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017E1726.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017E1726
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .908
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9017E1726"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           34.4%;  Score 755;  DB 7;  Length 908;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  805;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 84

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 144

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 204

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 264

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        265 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 324

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 384

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 444

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 504

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 564

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 624

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 684

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 744

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 804

Qy        846 GACAAGGAAGGCATTCCTCCTGACCA 871
              ||||||||||||||||||||||||||
Db        805 GACAAGGAAGGCATTCCTCCTGACCA 830


RESULT 123
CT001300
LOCUS       CT001300                922 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT001300 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017L1615 5',
            mRNA sequence.
ACCESSION   CT001300
VERSION     CT001300.1  GI:68295183
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
```

```
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017L1615.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017L1615
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES              Location/Qualifiers
     source           1. .922
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="RZPDp9017L1615"
                      /tissue_type="T-Lymphocytes"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /clone_lib="RZPD no.9017"
                      /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                      NotI; vector:
                      http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                      ; 1st strand cDNA was prepared from mRNA obtained from
                      human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                      GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                      Double-stranded cDNA was ligated to SalI adaptors,
                      digested with NotI and cloned into the NotI and SalI sites
                      of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           34.4%;  Score 755;  DB 7;  Length 922;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  805;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 111

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 171

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 231

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 291

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        292 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 351

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 411

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 471

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 531

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 591

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 651

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 771

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 831

Qy        846 GACAAGGAAGGCATTCCTCCTGACCA 871
              ||||||||||||||||||||||||||
Db        832 GACAAGGAAGGCATTCCTCCTGACCA 857


RESULT 124
CD105734
LOCUS       CD105734                859 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  AGENCOURT_13979152 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30367842 5', mRNA sequence.
ACCESSION   CD105734
VERSION     CD105734.1  GI:30758908
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM429  row: k  column: 19
            High quality sequence stop: 674.
FEATURES              Location/Qualifiers
     source           1. .859
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:30367842"
                      /tissue_type="Pituitary"
                      /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                      /clone_lib="NIH_MGC_179"
                      /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                      (destroyed); Site_2: NotI; Library is oligo-dT primed and
                      directionally cloned (EcoRV site is destroyed upon
                      cloning).  Average insert size 1.1 kb. Library was
                      constructed by (Invitrogen). Note: this is a NIH_MGC
                      Library."
ORIGIN
```

```
  Query Match             34.4%;  Score 754;  DB 6;  Length 859;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  754;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        282 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 60

Qy        342 GAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGC 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGC 120

Qy        402 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACC 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACC 180

Qy        462 CTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGT 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGT 240

Qy        522 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCC 300

Qy        582 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGAT 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGAT 360

Qy        642 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 420

Qy        702 AACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATC 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATC 480

Qy        762 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATC 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATC 540

Qy        822 GAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 600

Qy        882 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 660

Qy        942 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1001
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 720

Qy       1002 CTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 1035
              |||||||||||||||||||||||||||||||||
Db        721 CTGACTGGTAAGACCATCACCCTCGAAGTGGAGC 754


RESULT 125
CR984559
LOCUS       CR984559                909 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR984559 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016H044 5',
            mRNA sequence.
ACCESSION   CR984559
VERSION     CR984559.1  GI:68222893
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016H044.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016H044
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016H044"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. .
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.4%;  Score 753;  DB 7;  Length 909;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  803;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 115

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 175

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 235

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 295

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 355

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 415

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 475
```

```
Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 535

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 595

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 655

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        656 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 715

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        716 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 775

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        776 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 835

Qy        846 GACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||
Db        836 GACAAGGAAGGCATTCCTCCTGAC 859


RESULT 126
CT001212
LOCUS       CT001212                927 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT001212 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016B0427 5',
            mRNA sequence.
ACCESSION   CT001212
VERSION     CT001212.1  GI:68295095
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 927)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016B0427.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016B0427
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .927
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016B0427"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           34.4%;  Score 753;  DB 7;  Length 927;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  803;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 108

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 168

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 228

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 288

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        289 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 348

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 408

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 468

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 528

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 588

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 648

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 708

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 768

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 828

Qy        846 GACAAGGAAGGCATTCCTCCTGAC 869
              ||||||||||||||||||||||||
Db        829 GACAAGGAAGGCATTCCTCCTGAC 852
```

```
RESULT 127
CX867108
LOCUS       CX867108                753 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_15_D11.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7472327 5', mRNA sequence.
ACCESSION   CX867108
VERSION     CX867108.1  GI:58549664
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 753)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: HESC4_15_D11.b1_A037
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15777  row: h  column: 21
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 753.
FEATURES             Location/Qualifiers
     source          1. .753
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7472327"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             34.3%;  Score 752;  DB 8;  Length 753;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  752;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 61

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 121

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 181

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 241

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 301

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 361

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 421

Qy        748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 481

Qy        808 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTG 541

Qy        868 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 601

Qy        928 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 661

Qy        988 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 1047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCA 721

Qy       1048 TTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1079
              ||||||||||||||||||||||||||||||||
Db        722 TTGAGAATGTCAAGGCAAAGATCCAAGACAAG 753


RESULT 128
BQ888883
LOCUS       BQ888883                952 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8241462 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6180410 5', mRNA sequence.
ACCESSION   BQ888883
VERSION     BQ888883.1  GI:22280897
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 952)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13563  row: c  column: 03
```

```
            High quality sequence stop: 663.
FEATURES             Location/Qualifiers
     source          1. .952
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6180410"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           34.3%;  Score 752;  DB 5;  Length 952;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  752;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 767
              |||||||||||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAAATCTTCGTG 752


RESULT 129
CD244644
LOCUS       CD244644                 845 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14097252 NIH_MGC_180 Homo sapiens cDNA clone
            IMAGE:30377585 5', mRNA sequence.
ACCESSION   CD244644
VERSION     CD244644.1  GI:31005108
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 845)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM442  row: a  column: 18
            High quality sequence stop: 695.
FEATURES             Location/Qualifiers
     source          1. .845
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30377585"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_180"
                     /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI;
                     Site_2: EcoRV (destroyed); Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.68 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match           34.3%;  Score 751;  DB 6;  Length 845;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  751;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        293 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCC 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCC 60

Qy        353 CAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGA 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGA 120

Qy        413 CCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTA 180

Qy        473 CAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGAT 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGAT 240

Qy        533 CTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCAT 592
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCAT 300

Qy        593 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTT 652
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTT 360

Qy        653 GATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAA 712
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAA 420

Qy        713 AGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 772
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 480

Qy        773 ACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 832
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 540

Qy        833 AGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 892
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 600

Qy        893 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 952
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 660

Qy        953 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1012
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 720

Qy       1013 GACCATCACCCTCGAAGTGGAGCCGAGTGAC 1043
                 ||||||||||||||||||||||||||||||
Db        721 GACCATCACCCTCGAAGTGGAGCCGAGTGAC 751


RESULT 130
BQ881766
LOCUS       BQ881766                855 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8729983 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6339432
            5', mRNA sequence.
ACCESSION   BQ881766
VERSION     BQ881766.1  GI:22273774
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 855)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2534  row: e  column: 01
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .855
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6339432"
                     /tissue_type="neuroblastoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_47"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             34.3%;  Score 751;  DB 5;  Length 855;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  751;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 TCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACA 68
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACA 60

Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 120

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 180

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 240

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 300

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 360

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 420

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 480

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 540

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 600

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 660

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 720

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAA 759
                 ||||||||||||||||||||||||||||||
Db        721 CTGGTACTCCGTCTCAGAGGTGGGATGCAAA 751


RESULT 131
DR762082
LOCUS       DR762082                750 bp    mRNA    linear   EST 21-JUL-2005
DEFINITION  HESC4_136_C01.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7971556 5', mRNA sequence.
```

```
ACCESSION   DR762082
VERSION     DR762082.1  GI:71054782
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 750)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM 17088  row: f  column: 02
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 750.
FEATURES             Location/Qualifiers
     source          1. .750
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7971556"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN
```

Query Match 34.2%; Score 750; DB 8; Length 750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          21 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 60

Qy          81 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 120

Qy         141 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 180

Qy         201 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 240

Qy         261 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 300

Qy         321 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 360

Qy         381 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 420

Qy         441 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 480

Qy         501 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 540

Qy         561 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 600

Qy         621 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 660

Qy         681 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 720

Qy         741 CTCAGAGGTGGGATGCAAATCTTCGTGAAG 770
               |||||||||||||||||||||||||||||
Db         721 CTCAGAGGTGGGATGCAAATCTTCGTGAAG 750
```

```
RESULT 132
BM455235
LOCUS       BM455235                1114 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6405614 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500187
            5', mRNA sequence.
ACCESSION   BM455235
VERSION     BM455235.1  GI:18504275
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1114)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12134  row: l  column: 12
            High quality sequence stop: 692.
FEATURES             Location/Qualifiers
     source          1. .1114
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
                     /clone="IMAGE:5500187"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             34.2%;  Score 750;  DB 3;  Length 1114;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  750;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 65

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 125

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 185

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 245

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 305

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 365

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 425

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 485

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 545

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 605

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 665

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 725

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGG 781
              ||||||||||||||||||||||||||||||
Db        726 GATGCAAATCTTCGTGAAGACACTCACTGG 755


RESULT 133
CR985228
LOCUS       CR985228                910 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR985228 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017J032 5',
            mRNA sequence.
ACCESSION   CR985228
VERSION     CR985228.1  GI:68279113
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 910)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017J032.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017J032
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .910
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017J032"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.2%;  Score 749;  DB 7;  Length 910;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  749;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 60

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 120

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 180

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 240

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 300

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 360

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 420

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 480

Qy        729 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 540

Qy        789 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGAC 600

Qy        849 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 660

Qy        909 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 720

Qy        969 CTCAGAGGTGGGATGCAGATCTTCGTGAA 997
              |||||||||||||||||||||||||||||
Db        721 CTCAGAGGTGGGATGCAGATCTTCGTGAA 749


RESULT 134
CF131969
LOCUS       CF131969                768 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-aws-c-20-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30553699 5', mRNA sequence.
ACCESSION   CF131969
VERSION     CF131969.1  GI:33214759
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .768
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30553699"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

  Query Match             34.1%;  Score 748;  DB 6;  Length 768;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  748;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAAT 70
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAAT 60

Qy         71 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 120

Qy        131 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 180

Qy        191 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 240

Qy        251 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 300

Qy        311 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 360

Qy        371 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 420

Qy        431 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 480

Qy        491 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 540

Qy        551 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 600

Qy        611 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 660

Qy        671 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 720

Qy        731 GGTACTCCGTCTCAGAGGTGGGATGCAA 758
              ||||||||||||||||||||||||||||
Db        721 GGTACTCCGTCTCAGAGGTGGGATGCAA 748
```

```
RESULT 135
CF131877
LOCUS       CF131877                817 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-aws-a-21-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30553652 5', mRNA sequence.
ACCESSION   CF131877
VERSION     CF131877.1  GI:33214575
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 817)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .817
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30553652"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

  Query Match           34.1%;  Score 748;  DB 6;  Length 817;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  748;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAAT 70
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAAT 60

Qy         71 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 120

Qy        131 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 180

Qy        191 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 240

Qy        251 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 300

Qy        311 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 360

Qy        371 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 420

Qy        431 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 480

Qy        491 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 540

Qy        551 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 600

Qy        611 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 660

Qy        671 GGTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 720

Qy        731 GGTACTCCGTCTCAGAGGTGGGATGCAA 758
              ||||||||||||||||||||||||||||
Db        721 GGTACTCCGTCTCAGAGGTGGGATGCAA 748


RESULT 136
BQ233859
LOCUS       BQ233859                899 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7594233 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6021590
            5', mRNA sequence.
ACCESSION   BQ233859
VERSION     BQ233859.1  GI:20415259
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 899)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13227  row: i  column: 15
            High quality sequence stop: 586.
FEATURES             Location/Qualifiers
     source          1. .899
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

/clone="IMAGE:6021590"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 34.1%; Score 748; DB 3; Length 899;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 798; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        858 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACC 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACC 68

Qy        918 CTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGT 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGT 128

Qy        978 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCG 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCG 188

Qy       1038 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 248

Qy       1098 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 308

Qy       1158 AACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 AACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 368

Qy       1218 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATT 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATT 428

Qy       1278 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 488

Qy       1338 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 548

Qy       1398 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 608

Qy       1458 CTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 CTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAG 668

Qy       1518 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGG 728

Qy       1578 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTG 788

Qy       1638 CACCTGGTGCTCCGTCTTA 1656
              ||||||||||||||||||
Db        789 CACCTGGTGCTCCGTCTTA 807
```

RESULT 137
BM550155

```
LOCUS       BM550155                1004 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6558133 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5443377
            5', mRNA sequence.
ACCESSION   BM550155
VERSION     BM550155.1  GI:18786122
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1004)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1920  row: m  column: 10
            High quality sequence stop: 633.
FEATURES             Location/Qualifiers
     source          1. .1004
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5443377"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Library constructed by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 34.1%; Score 748; DB 3; Length 1004;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        276 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATC 60

Qy        336 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAG 120

Qy        396 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGG 180

Qy        456 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTC 240

Qy        516 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTG 300

Qy        576 GAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCT 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCT 360
```

```
Qy        636 CCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCT 695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCT 420

Qy        696 GACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATG 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATG 480

Qy        756 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGAC 540

Qy        816 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAG 600

Qy        876 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 660

Qy        936 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 720

Qy        996 AAGACCCTGACTGGTAAGACCATCACCC 1023
              ||||||||||||||||||||||||||||
Db        721 AAGACCCTGACTGGTAAGACCATCACCC 748
```

RESULT 138
BX358407
LOCUS BX358407 1074 bp mRNA linear EST 08-APR-2004
DEFINITION BX358407 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI037YD16 5-PRIME, mRNA sequence.
ACCESSION BX358407
VERSION BX358407.2 GI:46286658
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30368220.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6328.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI037DB08QP1&c=6328.r.
FEATURES Location/Qualifiers
source 1. .1074
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI037YD16"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

Query Match 34.1%; Score 748; DB 5; Length 1074;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 798; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         35 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 94

Qy        109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 154

Qy        169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 214

Qy        229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 274

Qy        289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        275 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 334

Qy        349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 394

Qy        409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 454

Qy        469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 514

Qy        529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 574

Qy        589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 634

Qy        649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        635 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 694

Qy        709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        695 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 754

Qy        769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        755 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 814

Qy        829 TCAAAGCAAAGATCCAAGA 847
              |||||||||||||||||||
Db        815 TCAAAGCAAAGATCCAAGA 833
```

RESULT 139
BQ050835
LOCUS BQ050835 995 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_6808984 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5784838 5', mRNA sequence.
ACCESSION BQ050835
VERSION BQ050835.1 GI:19810163
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 995)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12870  row: p  column: 23
            High quality sequence stop: 641.
FEATURES             Location/Qualifiers
     source          1. .995
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5784838"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match           34.1%;  Score 747;  DB 3;  Length 995;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  747;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        953 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 60

Qy       1013 GACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCA 120

Qy       1073 AGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 180

Qy       1133 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCT 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCT 240

Qy       1193 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCT 300

Qy       1253 CGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGG 360

Qy       1313 CATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCAC 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCAC 420

Qy       1373 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGG 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGG 480

Qy       1433 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCC 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCC 540

Qy       1493 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGA 600

Qy       1553 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTA 660

Qy       1613 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGAT 720

Qy       1673 CTTCGTGAAGACCCTGACTGGTAAGAC 1699
              |||||||||||||||||||||||||||
Db        721 CTTCGTGAAGACCCTGACTGGTAAGAC 747


RESULT 140
AW409943/c
LOCUS       AW409943                774 bp    mRNA    linear   EST 29-JUN-2000
DEFINITION  fh03c07.y1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961036 3',
            mRNA sequence.
ACCESSION   AW409943
VERSION     AW409943.1  GI:6935484
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Plate: LLCM55  row: F  column: 13
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2961036"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_17"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match           34.0%;  Score 746;  DB 1;  Length 774;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  746;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1447 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTG 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        774 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTG 715

Qy       1507 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGA 1566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        714 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGA 655
```

```
Qy       1567 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 1626
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        654 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 595

Qy       1627 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCC 1686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCC 535

Qy       1687 TGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGG 1746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 TGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGG 475

Qy       1747 CAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGA 1806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 CAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGA 415

Qy       1807 AACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 1866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 AACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 355

Qy       1867 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 1926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 295

Qy       1927 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 1986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 235

Qy       1987 ATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 2046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAG 175

Qy       2047 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGC 2106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 ATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGC 115

Qy       2107 GCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTT 2166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GCTTGAGGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTT 55

Qy       2167 CCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||
Db          54 CCTTTCAATAAAGTTGTTGCATTCCC 29
```

RESULT 141
CF552282

```
LOCUS       CF552282                878 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15608852 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30530405 5', mRNA sequence.
ACCESSION   CF552282
VERSION     CF552282.1  GI:34889116
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: NDAM619  row: i  column: 06
             High quality sequence stop: 652.
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30530405"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN
```

```
  Query Match             34.0%;  Score 746;  DB 6;  Length 878;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  796;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGANAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 720

Qy        748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 807
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        721 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGC 780

Qy        808 CCAGTGACACTATCGAG 824
              |||||||||||||||||
Db        781 CCAGTGACACTATCGAG 797


RESULT 142
BQ215503
LOCUS       BQ215503                884 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7511156 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6055534
            5', mRNA sequence.
ACCESSION   BQ215503
VERSION     BQ215503.1  GI:20396903
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13315  row: o  column: 23
            High quality sequence stop: 697.
FEATURES             Location/Qualifiers
     source          1. .884
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6055534"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             34.0%;  Score 746;  DB 3;  Length 884;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  746;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 GATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCT 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCT 60

Qy         79 TCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCG 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCG 120

Qy        139 AGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGA 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGA 180

Qy        199 TCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAG 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAG 240

Qy        259 AGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACAC 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACAC 300

Qy        319 TCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAG 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAG 360

Qy        379 CAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 420

Qy        439 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 480

Qy        499 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 540

Qy        559 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 600

Qy        619 ATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 660

Qy        679 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 720

Qy        739 GTCTCAGAGGTGGGATGCAAATCTTC 764
              |||||||||||||||||||||||||
Db        721 GTCTCAGAGGTGGGATGCAAATCTTC 746


RESULT 143
CT001459
LOCUS       CT001459                938 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT001459 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017H185 5',
            mRNA sequence.
ACCESSION   CT001459
VERSION     CT001459.1  GI:68295342
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 938)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017H185.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017H185
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .938
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017H185"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             34.0%;  Score 745;  DB 7;  Length 938;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  745;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          53 TGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCT 112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 TGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCT 100

Qy         113 CGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 CGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 160

Qy         173 CATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCAC 232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         161 CATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCAC 220

Qy         233 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGG 292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         221 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGG 280

Qy         293 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCC 352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         281 TGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCC 340

Qy         353 CAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGA 412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         341 CAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGA 400

Qy         413 CCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTA 472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         401 CCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTA 460

Qy         473 CAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGAT 532
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         461 CAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGAT 520

Qy         533 CTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCAT 592
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         521 CTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCAT 580

Qy         593 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTT 652
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         581 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTT 640

Qy         653 GATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAA 712
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         641 GATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAA 700

Qy         713 AGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         701 AGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 760

Qy         773 ACTCACTGGCAAGACCATCACCCTT 797
               |||||||||||||||||||||||||
Db         761 ACTCACTGGCAAGACCATCACCCTT 785
```

```
RESULT 144
CX786633
LOCUS       CX786633                744 bp    mRNA    linear   EST 02-MAR-2005
DEFINITION  HESC3_62_G06.g1_A036 NIH_MGC_260 Homo sapiens cDNA clone
            IMAGE:7481654 5', mRNA sequence.
ACCESSION   CX786633
VERSION     CX786633.1  GI:58303423
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15801  row: m  column: 12
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 744.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7481654"
                     /sex="male"
                     /tissue_type="human embryonic stem cells"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_260"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     SSEA3, SSEA4, Tra 1-60, Tra 1-81, CD9, Alk Phos, Oct4 and
                     Nanog expression; negative for SSEA1 expression. Passage
                     number 21. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_261). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             33.9%;  Score 744;  DB 8;  Length 744;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  744;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 60

Qy         846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 120
```

```
Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 180

Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 240

Qy        1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 300

Qy        1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 360

Qy        1146 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 1205
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGT 420

Qy        1206 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 1265
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCG 480

Qy        1266 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 1325
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAT 540

Qy        1326 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1385
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 600

Qy        1386 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1445
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 660

Qy        1446 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1505
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 720

Qy        1506 GAGAATGTCAAGGCAAAGATCCAA 1529
               |||||||||||||||||||||||||
Db         721 GAGAATGTCAAGGCAAAGATCCAA 744
```

RESULT 145
CF131596
LOCUS CF131596 753 bp mRNA linear EST 05-AUG-2003
DEFINITION UI-HF-FQ0-awr-g-07-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone IMAGE:30555318 5', mRNA sequence.
ACCESSION CF131596
VERSION CF131596.1 GI:33214011
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 753)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .753
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30555318"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

Query Match 33.9%; Score 744; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAA 758
                |||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAA 744


RESULT 146
CF132751
LOCUS       CF132751                  764 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-awq-o-07-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30558966 5', mRNA sequence.
ACCESSION   CF132751
VERSION     CF132751.1  GI:33216334
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 764)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .764
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30558966"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

  Query Match           33.9%;  Score 744;  DB 6;  Length 764;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  744;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCAA 758
              |||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCAA 744


RESULT 147
CX787889
LOCUS       CX787889                  766 bp    mRNA    linear   EST 02-MAR-2005
DEFINITION  HESC3_91_C04.g1_A036 NIH_MGC_260 Homo sapiens cDNA clone
            IMAGE:7484169 5', mRNA sequence.
ACCESSION   CX787889
VERSION     CX787889.1  GI:58304679
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 766)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM15808 row: f column: 07
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 766.

FEATURES Location/Qualifiers
source 1. .766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7484169"
/sex="male"
/tissue_type="human embryonic stem cells"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_260"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos. Cell line id and NIH Registry designation is BG01. Positive for SSEA3, SSEA4, Tra 1-60, Tra 1-81, CD9, Alk Phos, Oct4 and Nanog expression; negative for SSEA1 expression. Passage number 21. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_261). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 33.9%; Score 743; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          27 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 83

Qy          87 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 143

Qy         147 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 203

Qy         207 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 263

Qy         267 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         264 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 323

Qy         327 AAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATC 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         324 AAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATC 383

Qy         387 CAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 446
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 CAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 443

Qy         447 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 506
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 503

Qy         507 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 566
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 563

Qy         567 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         564 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 623

Qy         627 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 686
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         624 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 683

Qy         687 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 746
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         684 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 743

Qy         747 GGTGGGATGCAAATCTTCGTGAA 769
               |||||||||||||||||||||||
Db         744 GGTGGGATGCAAATCTTCGTGAA 766
```

RESULT 148
CF132418
LOCUS CF132418 768 bp mRNA linear EST 05-AUG-2003
DEFINITION UI-HF-FQ0-awo-k-10-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone IMAGE:30560025 5', mRNA sequence.
ACCESSION CF132418
VERSION CF132418.1 GI:33215665
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 768)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.

FEATURES Location/Qualifiers
source 1. .768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30560025"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 33.9%; Score 743; DB 6; Length 768;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCA 757
              |||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCA 743
```

RESULT 149
DR759783
LOCUS DR759783 782 bp mRNA linear EST 21-JUL-2005
DEFINITION HESC4_107_H02.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone IMAGE:7969109 5', mRNA sequence.
ACCESSION DR759783
VERSION DR759783.1 GI:71052483
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 782)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM 17081 row: p column: 03
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 782.
FEATURES Location/Qualifiers
source 1. .782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7969109"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 33.9%; Score 743; DB 8; Length 782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300
```

```
Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGCA 757
              |||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGCA 743
```

```
RESULT 150
CF132037
LOCUS       CF132037                813 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-aws-b-11-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30553666 5', mRNA sequence.
ACCESSION   CF132037
VERSION     CF132037.1  GI:33214894
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 813)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .813
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30553666"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN
```

```
  Query Match             33.9%;  Score 743;  DB 6;  Length 813;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  743;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGGATGCAA 758
              |||||||||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGGATGCAA 743
```

```
RESULT 151
CT001169
LOCUS       CT001169                912 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT001169 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017L0616 5',
            mRNA sequence.
ACCESSION   CT001169
VERSION     CT001169.1  GI:68295052
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 912)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017L0616.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017L0616
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .912
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017L0616"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             33.9%;  Score 743;  DB 7;  Length 912;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  793;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 111

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 171

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 231

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 291

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        292 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 351

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 411

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 471

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 531

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 591

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 651

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 771

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 831

Qy        846 GACAAGGAAGGCAT 859
              ||||||||||||||
Db        832 GACAAGGAAGGCAT 845


RESULT 152
BM462979
LOCUS       BM462979                1135 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6447335 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5537643
            5', mRNA sequence.
ACCESSION   BM462979
VERSION     BM462979.1  GI:18512019
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1135)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12229  row: e  column: 04
            High quality sequence stop: 708.
FEATURES             Location/Qualifiers
```

```
     source          1. .1135
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5537643"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.9%;  Score 743;  DB 3;  Length 1135;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  743;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 60

Qy         73 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 120

Qy        133 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 180

Qy        193 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 240

Qy        253 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 300

Qy        313 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 360

Qy        373 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 420

Qy        433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 480

Qy        493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 540

Qy        553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 600

Qy        613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 660

Qy        673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 720

Qy        733 TACTCCGTCTCAGAGGTGGGATG 755
              |||||||||||||||||||||||
Db        721 TACTCCGTCTCAGAGGTGGGATG 743


RESULT 153
BM999891/c
LOCUS       BM999891                760 bp    mRNA    linear   EST 17-JUN-2002
DEFINITION  UI-H-DP0-avs-c-01-0-UI.s1 NCI_CGAP_Fs1 Homo sapiens cDNA clone
            IMAGE:5883576 3', mRNA sequence.
ACCESSION   BM999891
VERSION     BM999891.1  GI:19724799
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Mary Hendrix
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5883576"
                     /tissue_type="Fibrosarcoma"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Fs1"
                     /note="Vector: pT7T3-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_Fs1 is
                     a cDNA library containing the following tissue(s):
                     Fibrosarcoma Cell line HT-1088 (ATCC number CCL-121). The
                     library was constructed according to Bonaldo, Lennon and
                     Soares, Genome Research, 6:791-806, 1996. First strand
                     cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site. Double stranded cDNA was ligated
                     to an EcoR I adaptor, digested with Not I, and cloned
                     directionally into pT7T3-Pac vector. The oligonucleotide
                     used to prime the synthesis of first-strand cDNA contains
                     a library tag sequence that is located between the Not I
                     site and the (dT)18 tail. The sequence tag for this
                     library is GTTCTACGAG.
                     TAG_TISSUE=fibrosarcoma
                     TAG_LIB=UI-H-DP0
                     TAG_SEQ=GTTCTACGAG"
ORIGIN

  Query Match             33.9%;  Score 742;  DB 3;  Length 760;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  742;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1451 GAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAA 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 GAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAA 701

Qy       1511 TGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTT 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTT 641

Qy       1571 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         640 TGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 581

Qy       1631 CACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         580 CACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 521
```

```
Qy       1691 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 TGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAA 461

Qy       1751 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 1810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACA 401

Qy       1811 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCT 341

Qy       1871 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 281

Qy       1931 CACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAA 1990
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 CACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAA 221

Qy       1991 GGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 2050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 161

Qy       2051 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTT 2110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTT 101

Qy       2111 GAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTT 2170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 GAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTT 41

Qy       2171 TCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||
Db          40 TCAATAAAGTTGTTGCATTCCC 19
```

RESULT 154
CF131371
LOCUS CF131371 780 bp mRNA linear EST 05-AUG-2003
DEFINITION UI-HF-FQ0-awp-j-09-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone IMAGE:30561152 5', mRNA sequence.
ACCESSION CF131371
VERSION CF131371.1 GI:33213558
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .780
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30561152"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

Query Match 33.9%; Score 742; DB 6; Length 780;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 742; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAGGTGGGATGC 756
```

```
                    ||||||||||||||||||||||
Db        721 CTCCGTCTCAGAGGTGGGATGC 742


RESULT 155
CF132239
LOCUS       CF132239                808 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-awr-h-22-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30555357 5', mRNA sequence.
ACCESSION   CF132239
VERSION     CF132239.1  GI:33215304
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 808)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .808
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30555357"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

  Query Match           33.9%;  Score 742;  DB 6;  Length 808;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  742;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 60

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 120

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 180

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 240

Qy        257 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 300

Qy        317 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 360

Qy        377 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 420

Qy        437 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 480

Qy        497 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 540

Qy        557 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 600

Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 660

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 720

Qy        737 CCGTCTCAGAGGTGGGATGCAA 758
              |||||||||||||||||||||||
Db        721 CCGTCTCAGAGGTGGGATGCAA 742


RESULT 156
BQ879520
LOCUS       BQ879520                857 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8118127 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6181248 5', mRNA sequence.
ACCESSION   BQ879520
VERSION     BQ879520.1  GI:22271528
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 857)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13565  row: f  column: 01
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .857
```

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6181248"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match             33.9%;  Score 742;  DB 5;  Length 857;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  742;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        507 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 60

Qy        567 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 120

Qy        627 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 180

Qy        687 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 240

Qy        747 GGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAG 806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAG 300

Qy        807 CCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCT 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCT 360

Qy        867 GACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGAC 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGAC 420

Qy        927 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 480

Qy        987 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACC 1046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACC 540

Qy       1047 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 1106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGG 600

Qy       1107 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 1166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 660

Qy       1167 AAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 1226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 720

Qy       1227 ACCCTGACTGGTAAGACCATCA 1248
              |||||||||||||||||||||
Db        721 ACCCTGACTGGTAAGACCATCA 742


RESULT 157
CF552062
LOCUS       CF552062                882 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15595077 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529906 5', mRNA sequence.
ACCESSION   CF552062
VERSION     CF552062.1  GI:34888896
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM618  row: d  column: 11
            High quality sequence start: 15
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529906"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match             33.9%;  Score 742;  DB 6;  Length 882;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  742;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        339 CTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAA 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 CTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAA 82

Qy        399 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 142

Qy        459 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 202

Qy        519 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 578
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 262

Qy        579 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCT 638
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         263 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCT 322

Qy         639 GATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 698
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         323 GATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGAC 382

Qy         699 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 758
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         383 TACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 442

Qy         759 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACT 818
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         443 ATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACT 502

Qy         819 ATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 878
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         503 ATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGG 562

Qy         879 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 938
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         563 TTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 622

Qy         939 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 998
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         623 AAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAG 682

Qy         999 ACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 1058
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         683 ACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTC 742

Qy        1059 AAGGCAAAGATCCAAGACAAGG 1080
               ||||||||||||||||||||||
Db         743 AAGGCAAAGATCCAAGACAAGG 764
```

```
RESULT 158
CF552518
LOCUS       CF552518                 882 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15595472 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529355 5', mRNA sequence.
ACCESSION   CF552518
VERSION     CF552518.1  GI:34889352
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM616  row: m  column: 12
            High quality sequence start: 4
            High quality sequence stop: 639.
FEATURES             Location/Qualifiers
     source          1..882
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529355"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match             33.9%;  Score 742;  DB 6;  Length 882;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  742;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          29 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 88

Qy          76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          89 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 148

Qy         136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         149 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 208

Qy         196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         209 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 268

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         269 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 328

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         329 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 388

Qy         376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         389 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 448

Qy         436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         449 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 508

Qy         496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         509 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 568

Qy         556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         569 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 628

Qy         616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         629 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 688

Qy         676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         689 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 748

Qy         736 TCCGTCTCAGAGGTGGGATGCA 757
               ||||||||||||||||||||||
Db         749 TCCGTCTCAGAGGTGGGATGCA 770
```

```
RESULT 159
BQ722449
LOCUS       BQ722449                 920 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8214978 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6187445 5', mRNA sequence.
ACCESSION   BQ722449
```

```
VERSION     BQ722449.1  GI:21861346
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 920)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13581  row: h  column: 06
            High quality sequence stop: 612.
FEATURES             Location/Qualifiers
     source          1. .920
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6187445"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.9 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine); available through Life
                     Technologies."
ORIGIN

  Query Match             33.8%;  Score 741;  DB 5;  Length 920;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  791;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 60

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 120

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 180

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 240

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 300

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 360

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 420

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 480

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAA 540

Qy        906 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTC 600

Qy        966 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTC 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        601 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 660

Qy       1026 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGC 720

Qy       1086 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACC 780

Qy       1146 CTGTCTGACTAC 1157
              ||||||||||||
Db        781 CTGTCTGACTAC 792


RESULT 160
BQ934306
LOCUS       BQ934306                958 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8821966 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6203369 5', mRNA sequence.
ACCESSION   BQ934306
VERSION     BQ934306.1  GI:22349689
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 958)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13622  row: o  column: 18
            High quality sequence stop: 632.
FEATURES             Location/Qualifiers
     source          1. .958
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6203369"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sciatic_nerve"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
```

NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 33.8%; Score 740; DB 5; Length 958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 83

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 143

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 203

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 263

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 323

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 383

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 443

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 503

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 563

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 623

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 683

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 743

Qy        742 TCAGAGGTGGGATGCAAATC 761
              ||||||||||||||||||||
Db        744 TCAGAGGTGGGATGCAAATC 763
```

RESULT 161
CR983401
LOCUS CR983401 893 bp mRNA linear EST 24-JUN-2005
DEFINITION CR983401 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017J102 5', mRNA sequence.
ACCESSION CR983401
VERSION CR983401.1 GI:68221735
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 893)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9017J102.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017J102
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9017J102"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 33.7%; Score 738; DB 7; Length 893;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 788; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 111

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 171

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 231

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 291

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
```

```
Db        292 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 351

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 411

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 471

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 531

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 591

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 651

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 771

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 831

Qy        846 GACAAGGAA 854
              |||||||||
Db        832 GACAAGGAA 840


RESULT 162
BQ226817
LOCUS       BQ226817                 897 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7553147 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6057911
            5', mRNA sequence.
ACCESSION   BQ226817
VERSION     BQ226817.1  GI:20408217
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13322  row: b  column: 24
            High quality sequence stop: 696.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6057911"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.  Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match             33.7%;  Score 738;  DB 3;  Length 897;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  738;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        670 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 AGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACC 62

Qy        730 TGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCA 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 TGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCA 122

Qy        790 TCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACA 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACA 182

Qy        850 AGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATG 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATG 242

Qy        910 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTC 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTC 302

Qy        970 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAG 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAG 362

Qy       1030 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 422

Qy       1090 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGT 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGT 482

Qy       1150 CTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGA 542

Qy       1210 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTG 1269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTG 602

Qy       1270 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGC 1329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGC 662

Qy       1330 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 1389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACA 722

Qy       1390 TCCAGAAAGAGTCCACCC 1407
              ||||||||||||||||||
Db        723 TCCAGAAAGAGTCCACCC 740


RESULT 163
AL536994
LOCUS       AL536994                 737 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL536994 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF017YB07 5-PRIME, mRNA sequence.
ACCESSION   AL536994
VERSION     AL536994.3  GI:45712790
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 737)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:31261618.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF017CA04QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .737
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF017YB07"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             33.6%;  Score 737;  DB 1;  Length 737;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  737;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 TCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACA 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACA 60

Qy         69 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGT 120

Qy        129 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAG 180

Qy        189 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 240

Qy        249 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTC 300

Qy        309 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAG 360

Qy        369 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATC 420

Qy        429 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        421 TTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 480

Qy        489 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTG 540

Qy        549 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 600

Qy        609 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAA 660

Qy        669 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCAC 720

Qy        729 CTGGTACTCCGTCTCAG 745
              |||||||||||||||||
Db        721 CTGGTACTCCGTCTCAG 737


RESULT 164
CD519152
LOCUS       CD519152                859 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14372818 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30400041 5', mRNA sequence.
ACCESSION   CD519152
VERSION     CD519152.1  GI:31450870
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM487  row: i  column: 10
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1. .859
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30400041"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             33.6%;  Score 737;  DB 6;  Length 859;
  Best Local Similarity   100.0%;  Pred. No. 0;
```

```
  Matches  737;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 86

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 146

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 206

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 266

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 326

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 386

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 446

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 506

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 566

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 626

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 686

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 746

Qy        736 TCCGTCTCAGAGGTGGG 752
              |||||||||||||||||
Db        747 TCCGTCTCAGAGGTGGG 763


RESULT 165
BU182397
LOCUS       BU182397                945 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7943331 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6161609
            5', mRNA sequence.
ACCESSION   BU182397
VERSION     BU182397.1  GI:22696381
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13514  row: c  column: 18
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .945
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6161609"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.6%;  Score 737;  DB 5;  Length 945;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  737;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660
```

```
Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGAGGTGGG 752
              |||||||||||||||||
Db        721 TCCGTCTCAGAGGTGGG 737


RESULT 166
DR761274
LOCUS       DR761274                734 bp    mRNA    linear   EST 21-JUL-2005
DEFINITION  HESC4_126_E05.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7970868 5', mRNA sequence.
ACCESSION   DR761274
VERSION     DR761274.1  GI:71053974
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM 17086  row: i  column: 10
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 734.
FEATURES              Location/Qualifiers
     source           1. .734
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:7970868"
                      /sex="male"
                      /tissue_type="embryonic stem"
                      /cell_type="human embryonic stem cells"
                      /cell_line="BG01"
                      /lab_host="DH10B-T1 phage-resistant E. coli"
                      /clone_lib="NIH_MGC_262"
                      /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                      RNA obtained from human embryonic stem cells isolated from
                      the inner cell mass of blastocyst stage embryos and
                      differentiated to an early neural progenitor cell type.
                      Cell line id and NIH Registry designation is BG01.
                      Positive for Nestin and Musashi expression. Passage number
                      18. cDNA primed using oligo-dT primer:
                      5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                      the EcoRV/NotI sites of pExpress-1. This primary library
                      is non-normalized (normalized primary library is
                      NIH_MGC_259). It was constructed by Express Genomics
                      (Frederick, MD). Sequence ends have been trimmed to
                      exclude vector and regions below Phred quality 16. Note:
                      this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             33.5%;  Score 734;  DB 8;  Length 734;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  734;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        984 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGAC 60

Qy       1044 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAG 120

Qy       1104 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATC 180

Qy       1164 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 240

Qy       1224 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 300

Qy       1284 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 360

Qy       1344 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 420

Qy       1404 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 480

Qy       1464 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 540

Qy       1524 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 600

Qy       1584 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 660

Qy       1644 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 720

Qy       1704 ACTCTCGAAGTGGA 1717
              ||||||||||||||
Db        721 ACTCTCGAAGTGGA 734


RESULT 167
CD517713
LOCUS       CD517713                888 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14372955 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30399928 5', mRNA sequence.
ACCESSION   CD517713
VERSION     CD517713.1  GI:31449431
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
```

```
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Dr. Michael Brownstein
           cDNA Library Preparation: Invitrogen Corp
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: NDAM487  row: d  column: 17.
          High quality sequence start: 5
          High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30399928"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             33.5%;  Score 734;  DB 6;  Length 888;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  784;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        471 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 87

Qy        531 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 147

Qy        591 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGG 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGG 207

Qy        651 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAG 267

Qy        711 AAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 AAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 327

Qy        771 ACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTC 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 ACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTC 387

Qy        831 AAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 447

Qy        891 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 950
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 507

Qy        951 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 1010
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 567

Qy       1011 AAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 AAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 627

Qy       1071 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         628 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 687

Qy       1131 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTG 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         688 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTG 747

Qy       1191 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1250
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         748 CTCCGTCTCAGAAGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 807

Qy       1251 CTCGA 1255
              |||||
Db         808 CTCGA 812


RESULT 168
CR978764
LOCUS       CR978764                915 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR978764 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016K2026 5',
            mRNA sequence.
ACCESSION   CR978764
VERSION     CR978764.1  GI:68217098
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016K2026.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016K2026
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016K2026"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. .
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN
```

```
  Query Match           33.5%;  Score 734;  DB 7;  Length 915;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  784;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        237 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGG 73

Qy        297 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGT 356
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         74 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGT 133

Qy        357 GACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAG 193

Qy        417 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 253

Qy        477 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 313

Qy        537 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAG 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAG 373

Qy        597 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATC 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATC 433

Qy        657 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAG 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAG 493

Qy        717 TCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 TCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 553

Qy        777 ACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 ACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCA 613

Qy        837 AAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 AAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 673

Qy        897 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 733

Qy        957 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 793

Qy       1017 ATCAC 1021
              |||||
Db        794 ATCAC 798


RESULT 169
BU146473
LOCUS       BU146473                982 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8229557 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6183767 5', mRNA sequence.
ACCESSION   BU146473
VERSION     BU146473.1  GI:22660005
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 982)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13571  row: n  column: 24
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .982
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6183767"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match           33.5%;  Score 734;  DB 5;  Length 982;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  734;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 85

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 145

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 205

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 265

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 325

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 385

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 445

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 505

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 565

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 625

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 685

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 745

Qy        735 CTCCGTCTCAGAGG 748
                ||||||||||||||
Db        746 CTCCGTCTCAGAGG 759


RESULT 170
CR991559
LOCUS       CR991559                733 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR991559 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016O1424 5',
            mRNA sequence.
ACCESSION   CR991559
VERSION     CR991559.1  GI:68285444
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016O1424.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016O1424
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016O1424"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           33.4%;  Score 733;  DB 7;  Length 733;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  733;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCAGAG 747
                |||||||||||||
Db        721 CTCCGTCTCAGAG 733


RESULT 171
BQ948667
LOCUS       BQ948667                917 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8854272 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6473633
            5', mRNA sequence.
ACCESSION   BQ948667
VERSION     BQ948667.1  GI:22364145
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14009  row: h  column: 18
            High quality sequence stop: 673.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6473633"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match            33.4%;  Score 733;  DB 5;  Length 917;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  733;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 69

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 129

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 189

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 249

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 309

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 369

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 429

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 489

Qy        504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 549

Qy        564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 609

Qy        624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 669

Qy        684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 729

Qy        744 AGAGGTGGGATGC 756
              |||||||||||||
Db        730 AGAGGTGGGATGC 742


RESULT 172
AL541647
LOCUS       AL541647                1074 bp    mRNA    linear   EST 24-MAR-2004
DEFINITION  AL541647 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE003YJ22
            5-PRIME, mRNA sequence.
ACCESSION   AL541647
VERSION     AL541647.3  GI:45717228
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1074)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:30546021.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DE003DE11QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1074
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE003YJ22"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match            33.4%;  Score 733;  DB 1;  Length 1074;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches  833;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 78

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 138
```

```
Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 198

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 258

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 318

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        319 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 378

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 438

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 498

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 558

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 618

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        619 TCCTCCTGATCAGCAGAGGTTGATCTTTKCCGGAAAACAGCTGGAAGATGGTCGTACCCT 678

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 738

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAG 798

Qy        812 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCT 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 TGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCT 853
```

```
RESULT 173
BU165709
LOCUS       BU165709                902 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7941040 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6155198
            5', mRNA sequence.
ACCESSION   BU165709
VERSION     BU165709.1  GI:22679650
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 902)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13497  row: h  column: 15
            High quality sequence stop: 573.
FEATURES             Location/Qualifiers
     source          1. .902
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6155198"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN
```

```
  Query Match            33.4%;  Score 732;  DB 5;  Length 902;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 71

Qy        738 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 131

Qy        798 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGC 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 GAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGC 191

Qy        858 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACC 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACC 251

Qy        918 CTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGT 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGT 311

Qy        978 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCG 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCG 371

Qy       1038 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 431

Qy       1098 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTAC 491

Qy       1158 AACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 AACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATC 551

Qy       1218 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATT 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATT 611

Qy       1278 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTG 671

Qy       1338 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAG 731

Qy       1398 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        732 GAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACC 791

Qy       1458 CTG 1460
              |||
Db        792 CTG 794


RESULT 174
CR977715
LOCUS       CR977715                925 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR977715 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016H2010 5',
            mRNA sequence.
ACCESSION   CR977715
VERSION     CR977715.1  GI:68216045
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016H2010.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016H2010
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016H2010"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             33.4%;  Score 732;  DB 7;  Length 925;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  782;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          49 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 108

Qy         126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         109 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 168

Qy         186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         169 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 228

Qy         246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 288

Qy         306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
               |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         289 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 348

Qy         366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         349 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 408

Qy         426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         409 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 468

Qy         486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         469 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 528

Qy         546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         529 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 588

Qy         606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         589 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 648

Qy         666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         649 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 708

Qy         726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         709 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 768

Qy         786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         769 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 828

Qy         846 GAC 848
               |||
Db         829 GAC 831


RESULT 175
CX873035
LOCUS       CX873035                731 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_76_A08.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7487538 5', mRNA sequence.
ACCESSION   CX873035
VERSION     CX873035.1  GI:58556209
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 731)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
```

```
           Tissue Procurement: BresaGen, Inc.
            cDNA Library Preparation: Express Genomics, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
           University of Georgia
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM15817  row: b  column: 16
           Seq primer: JENREV (CAGGAAACAGCTATGACC)
           High quality sequence stop: 731.
FEATURES             Location/Qualifiers
     source          1. .731
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7487538"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             33.3%;  Score 731;  DB 8;  Length 731;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  731;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 60

Qy         73 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 120

Qy        133 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 180

Qy        193 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 240

Qy        253 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 300

Qy        313 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 360

Qy        373 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 420

Qy        433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 480

Qy        493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 540

Qy        553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 600

Qy        613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 660

Qy        673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 720

Qy        733 TACTCCGTCTC 743
              ||||||||||||
Db        721 TACTCCGTCTC 731


RESULT 176
CN415690
LOCUS       CN415690                738 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600263767 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN415690
VERSION     CN415690.1  GI:47403284
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 738   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."
ORIGIN

  Query Match             33.3%;  Score 731;  DB 7;  Length 738;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  731;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 AGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGC 67
```

```
Qy        73 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        68 AGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACA 127

Qy       133 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       128 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGA 187

Qy       193 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       188 GGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 247

Qy       253 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 AGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 307

Qy       313 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 AGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACG 367

Qy       373 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 TCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTG 427

Qy       433 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       428 CCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTA 487

Qy       493 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 552
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       488 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 547

Qy       553 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 612
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       548 GTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGA 607

Qy       613 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       608 TCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGC 667

Qy       673 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       668 TGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGG 727

Qy       733 TACTCCGTCTC 743
             |||||||||||
Db       728 TACTCCGTCTC 738
```

```
RESULT 177
CD517104
LOCUS       CD517104                854 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14375178 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30395536 5', mRNA sequence.
ACCESSION   CD517104
VERSION     CD517104.1  GI:31448822
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 854)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM475  row: m  column: 17
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1. .854
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30395536"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 33.3%; Score 731; DB 6; Length 854;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 731; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        18 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 77
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 60

Qy        78 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 137
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 120

Qy       138 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 180

Qy       198 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 240

Qy       258 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 300

Qy       318 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 360

Qy       378 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 420

Qy       438 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 480

Qy       498 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 540

Qy       558 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 600

Qy       618 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 660

Qy       678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 720

Qy        738 CGTCTCAGAGG 748
                ||||||||||||
Db        721 CGTCTCAGAGG 731


RESULT 178
BQ224093
LOCUS       BQ224093                877 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7518547 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6056043
            5', mRNA sequence.
ACCESSION   BQ224093
VERSION     BQ224093.1  GI:20405493
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13317  row: e  column: 04
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1. .877
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6056043"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.3%;  Score 731;  DB 3;  Length 877;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  731;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCAGA 746
              ||||||||||||
Db        721 TCCGTCTCAGA 731


RESULT 179
CF552082
LOCUS       CF552082                925 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15595463 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529883 5', mRNA sequence.
ACCESSION   CF552082
VERSION     CF552082.1  GI:34888916
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM618  row: c  column: 12
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529883"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
```

```
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match           33.3%;  Score 731;  DB 6;  Length 925;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  731;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 TCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAG 87

Qy         87 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 ACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTC 147

Qy        147 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCT 207

Qy        207 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACC 267

Qy        267 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGC 327

Qy        327 AAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATC 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATC 387

Qy        387 CAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTG 447

Qy        447 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTG 507

Qy        507 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACC 567

Qy        567 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 CTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAA 627

Qy        627 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGT 687

Qy        687 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 747

Qy        747 GGTGGGATGCA 757
              ||||||||||||
Db        748 GGTGGGATGCA 758


RESULT 180
CD557608
LOCUS       CD557608                936 bp    mRNA    linear   EST 11-JUN-2003
DEFINITION  AGENCOURT_14413793 NIH_MGC_180 Homo sapiens cDNA clone
            IMAGE:30387377 5', mRNA sequence.
ACCESSION   CD557608
VERSION     CD557608.1  GI:31583676
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 936)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM457  row: i  column: 18
            High quality sequence stop: 616.
FEATURES             Location/Qualifiers
     source          1. .936
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30387377"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_180"
                     /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI;
                     Site_2: EcoRV (destroyed); Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.68 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match           33.3%;  Score 731;  DB 6;  Length 936;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  731;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 67

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 127

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 187

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 247

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 307

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 367

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 427

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        428 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 487

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 547

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 607

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 667

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 727

Qy        742 TCAGAGGTGGG 752
              ||||||||||| 
Db        728 TCAGAGGTGGG 738


RESULT 181
CT000208
LOCUS       CT000208                739 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT000208 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017M1720 5',
            mRNA sequence.
ACCESSION   CT000208
VERSION     CT000208.1  GI:68294091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017M1720.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017M1720
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017M1720"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             33.3%;  Score 730;  DB 7;  Length 739;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  730;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCA 744
              ||||||||||| 
Db        721 CTCCGTCTCA 730


RESULT 182
CD245746
LOCUS       CD245746                834 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14098110 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30376822 5', mRNA sequence.
ACCESSION   CD245746
VERSION     CD245746.1  GI:31006210
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM440  row: a  column: 23
            High quality sequence stop: 687.
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30376822"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           33.3%;  Score 730;  DB 6;  Length 834;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  730;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 82
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGT 60

Qy         83 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAA 120

Qy        143 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTT 180

Qy        203 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 240

Qy        263 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCAC 300

Qy        323 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAA 360

Qy        383 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 420

Qy        443 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 480

Qy        503 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        481 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 540

Qy        563 CACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAA 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAA 600

Qy        623 GGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGG 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGG 660

Qy        683 TCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCT 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCT 720

Qy        743 CAGAGGTGGG 752
              ||||||||||
Db        721 CAGAGGTGGG 730


RESULT 183
CD517028
LOCUS       CD517028                890 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14374773 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30407392 5', mRNA sequence.
ACCESSION   CD517028
VERSION     CD517028.1  GI:31448746
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 890)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM488  row: k  column: 17
            High quality sequence start: 8
            High quality sequence stop: 657.
FEATURES             Location/Qualifiers
     source          1. .890
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30407392"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           33.3%;  Score 730;  DB 6;  Length 890;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  730;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        964 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 1023
```

```
Db         29 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 88

Qy       1024 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1083

Db         89 TCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 148

Qy       1084 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 1143

Db        149 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 208

Qy       1144 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 1203

Db        209 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAG 268

Qy       1204 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 1263

Db        269 GTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGC 328

Qy       1264 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 1323

Db        329 CGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTG 388

Qy       1324 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 1383

Db        389 ATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACT 448

Qy       1384 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 1443

Db        449 ACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGA 508

Qy       1444 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 1503

Db        509 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCA 568

Qy       1504 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 1563

Db        569 TTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGT 628

Qy       1564 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 1623

Db        629 TGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGA 688

Qy       1624 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 1683

Db        689 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGA 748

Qy       1684 CCCTGACTGG 1693

Db        749 CCCTGACTGG 758


RESULT 184
CT003566
LOCUS       CT003566                908 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT003566 RZPD no.9017 Homo sapiens cDNA clone RZPDp901701825 5',
            mRNA sequence.
ACCESSION   CT003566
VERSION     CT003566.1  GI:68297449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp901701825.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp901701825
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .908
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp901701825"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. .
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             33.3%;  Score 730;  DB 7;  Length 908;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  730;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74

Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134

Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194

Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254

Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314

Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374

Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434

Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494

Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
```

```
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTCA 744
              ||||||||||
Db        721 CTCCGTCTCA 730


RESULT 185
BM563395
LOCUS       BM563395                970 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6589307 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5441640
            5', mRNA sequence.
ACCESSION   BM563395
VERSION     BM563395.1  GI:18810281
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1916  row: e  column: 01
            High quality sequence stop: 703.
FEATURES             Location/Qualifiers
     source          1. .970
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5441640"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             33.3%;  Score 730;  DB 3;  Length 970;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  780;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 60

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 120

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 180

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTATAACATCCAGAA 240

Qy        257 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 300

Qy        317 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 360

Qy        377 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 420

Qy        437 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 480

Qy        497 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 540

Qy        557 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 600

Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 660

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 720

Qy        737 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 780

Qy        797 T 797
              |
Db        781 T 781


RESULT 186
CX164520
LOCUS       CX164520                729 bp    mRNA    linear   EST 23-DEC-2004
DEFINITION  HESC2_17_F10.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7468173 5', mRNA sequence.
ACCESSION   CX164520
VERSION     CX164520.1  GI:56794600
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
```

```
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: BresaGen, Inc.
           cDNA Library Preparation: Express Genomics, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
          University of Georgia
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM15766  row: k  column: 19
          Seq primer: JENREV (CAGGAAACAGCTATGACC)
          High quality sequence stop: 729.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7468173"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
                     Genomics (Frederick, MD). Sequence ends have been trimmed
                     to exclude vector and regions below Phred quality 16.
                     Three-prime sequences are presented as their reverse
                     complement and have been trimmed to exclude polyA. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             33.3%;  Score 729;  DB 8;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTC 743
              |||||||||
Db        721 CTCCGTCTC 729


RESULT 187
CX873731
LOCUS       CX873731                729 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_84_D10.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7488454 5', mRNA sequence.
ACCESSION   CX873731
VERSION     CX873731.1  GI:58556905
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15819  row: h  column: 20
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 729.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7488454"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
```

/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 33.3%; Score 729; DB 8; Length 729;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy      75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy     135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy     195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy     255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy     315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy     375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy     435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy     495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy     555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy     615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy     675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy     735 CTCCGTCTC 743
           |||||||||
Db     721 CTCCGTCTC 729
```

RESULT 188
DN601704
LOCUS DN601704 729 bp mRNA linear EST 17-MAR-2005
DEFINITION HESC3_71_A04.g2_A036 NIH_MGC_260 Homo sapiens cDNA clone IMAGE:7482153 5', mRNA sequence.
ACCESSION DN601704
VERSION DN601704.1 GI:61644343
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM15803 row: b column: 07
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 729.
FEATURES Location/Qualifiers
source 1. .729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7482153"
/sex="male"
/tissue_type="human embryonic stem cells"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_260"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos. Cell line id and NIH Registry designation is BG01. Positive for SSEA3, SSEA4, Tra 1-60, Tra 1-81, CD9, Alk Phos, Oct4 and Nanog expression; negative for SSEA1 expression. Passage number 21. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_261). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 33.3%; Score 729; DB 8; Length 729;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy      88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120
```

```
Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 720

Qy        748 GTGGGATGC 756
              |||||||||
Db        721 GTGGGATGC 729
```

```
RESULT 189
DR759749
LOCUS       DR759749                729 bp    mRNA    linear   EST 21-JUL-2005
DEFINITION  HESC4_107_D12.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7968937 5', mRNA sequence.
ACCESSION   DR759749
VERSION     DR759749.1  GI:71052449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM 17081  row: h  column: 23
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 729.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7968937"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN
```

```
  Query Match             33.3%;  Score 729;  DB 8;  Length 729;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540
```

```
Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTC 743
              |||||||||
Db        721 CTCCGTCTC 729
```

RESULT 190
DR763210
LOCUS DR763210 729 bp mRNA linear EST 21-JUL-2005
DEFINITION HESC4_149_H05.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone IMAGE:7973315 5', mRNA sequence.
ACCESSION DR763210
VERSION DR763210.1 GI:71055910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 729)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM 17092 row: o column: 09
Seq primer: JENREV (CAGGAAACAGCTATGACC) .
High quality sequence stop: 729.
FEATURES Location/Qualifiers
source 1. .729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7973315"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."
ORIGIN

Query Match 33.3%; Score 729; DB 8; Length 729;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 720

Qy        735 CTCCGTCTC 743
              |||||||||
Db        721 CTCCGTCTC 729
```

RESULT 191
AL603620/c
LOCUS AL603620 780 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp434K0435_s1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434K0435 3', mRNA sequence.
ACCESSION AL603620
VERSION AL603620.1 GI:15167126
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 780)
  AUTHORS   Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Poustka, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            r1 sequence also available.
            This clone (DKFZp434K0435) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .780
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434K0435"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="434 (synonym: htes3)"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
ORIGIN

  Query Match             33.3%;  Score 729;  DB 1;  Length 780;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  779;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1406 CCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1465
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 721

Qy       1466 TAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGAT 1525
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        720 TAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGAT 661

Qy       1526 CCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 1585
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCT 601

Qy       1586 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 1645
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGT 541

Qy       1646 GCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 1705
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCAC 481

Qy       1706 TCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 1765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGA 421

Qy       1766 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 1825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACG 361

Qy       1826 CACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAG 1885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 CACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAG 301

Qy       1886 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 1945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 AGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGA 241

Qy       1946 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 2005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181

Qy       2006 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 2065
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGA 121

Qy       2066 CTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTA 2125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 CTACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTA 61

Qy       2126 AGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTG 2185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          60 AGTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTG 1


RESULT 192
BQ212090
LOCUS       BQ212090                879 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7600929 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6061365
            5', mRNA sequence.
ACCESSION   BQ212090
VERSION     BQ212090.1  GI:20391987
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13331  row: b  column: 22
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6061365"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.3%;  Score 729;  DB 3;  Length 879;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCGTCTCA 744
                    |||||||||
Db        721 TCCGTCTCA 729


RESULT 193
BQ227353
LOCUS       BQ227353                 893 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7418114 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:5772834
            5', mRNA sequence.
ACCESSION   BQ227353
VERSION     BQ227353.1  GI:20408753
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 893)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12839  row: l  column: 19
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1. .893
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5772834"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
ORIGIN

  Query Match             33.3%;  Score 729;  DB 3;  Length 893;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1148 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 1207
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 69

Qy       1208 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAG 1267
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAG 129

Qy       1268 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCA 1327
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCA 189

Qy       1328 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1387
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 249

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 309

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 369

Qy       1508 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 1567
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGAT 429

Qy       1568 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1627
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 489

Qy       1628 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1687
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 549

Qy       1688 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1747
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 609

Qy       1748 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1807
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 669

Qy       1808 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1867
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 729

Qy       1868 CCTGGTGCT 1876
                |||||||||
Db        730 CCTGGTGCT 738


RESULT 194
BM909385
```

```
LOCUS       BM909385               1053 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6640906 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5433955
            5', mRNA sequence.
ACCESSION   BM909385
VERSION     BM909385.1  GI:19359764
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1053)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1910  row: d  column: 20
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1. .1053
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5433955"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_99"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN
```

```
  Query Match             33.3%;  Score 729;  DB 3;  Length 1053;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         18 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATC 60

Qy         78 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATC 120

Qy        138 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTG 180

Qy        198 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 240

Qy        258 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 300

Qy        318 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAA 360

Qy        378 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 420

Qy        438 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 480

Qy        498 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 540

Qy        558 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAA 600

Qy        618 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 660

Qy        678 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTC 720

Qy        738 CGTCTCAGA 746
              |||||||||
Db        721 CGTCTCAGA 729
```

```
RESULT 195
CR989749
LOCUS       CR989749               865 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR989749 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017G1026 5',
            mRNA sequence.
ACCESSION   CR989749
VERSION     CR989749.1  GI:68283634
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 865)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017G1026.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017G1026
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .865
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017G1026"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
```

/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 33.2%; Score 728; DB 7; Length 865;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 828; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        592 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGT 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGT 84

Qy        652 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 144

Qy        712 AAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 AAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 204

Qy        772 CACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCA 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 CACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCA 264

Qy        832 AAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 AAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 324

Qy        892 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 384

Qy        952 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 444

Qy       1012 AGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 1071
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 AGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCC 504

Qy       1072 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 AAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 564

Qy       1132 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGC 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGC 624

Qy       1192 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTC 684

Qy       1252 TCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 TCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAG 744

Qy       1312 GCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 1371
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 GCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCA 804

Qy       1372 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 1421
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTC 854
```

RESULT 196
CX868946
LOCUS CX868946 727 bp mRNA linear EST 03-FEB-2005
DEFINITION HESC4_31_E01.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone IMAGE:7473891 5', mRNA sequence.
ACCESSION CX868946
VERSION CX868946.1 GI:58552120
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 727)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM15781 row: j column: 01
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 727.
FEATURES Location/Qualifiers
source 1. .727
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7473891"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 33.2%; Score 727; DB 8; Length 727;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120
```

```
Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 720

Qy        748 GTGGGAT 754
              |||||||
Db        721 GTGGGAT 727
```

```
RESULT 197
BU178242
LOCUS       BU178242                857 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7966425 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6165212
            5', mRNA sequence.
ACCESSION   BU178242
VERSION     BU178242.1  GI:22692226
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 857)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13523  row: i  column: 21
```

```
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .857
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6165212"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN
```

```
  Query Match             33.2%;  Score 727;  DB 5;  Length 857;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  827;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        718 CCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 71

Qy        778 CTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAA 131

Qy        838 AGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 AGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 191

Qy        898 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 251

Qy        958 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 311

Qy       1018 TCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACA 371

Qy       1078 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 AGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 431

Qy       1138 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTC 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTC 491

Qy       1198 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGG 551

Qy       1258 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 TGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCC 611

Qy       1318 CTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 CTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGT 671

Qy       1378 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGA 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 CTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTANAGGTGGGA 731

Qy       1438 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTG 1497
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        732 TGCAGATCTTCGTGAAGACCCTGACTGGTAAGAACATCACTCTCGAAGTGGAGCCGAGTG 791

Qy       1498 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1546
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db          792 ACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 840

RESULT 198
CR984951
LOCUS       CR984951                 861 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR984951 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017L034 5',
            mRNA sequence.
ACCESSION   CR984951
VERSION     CR984951.1  GI:68223285
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 861)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017L034.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017L034
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .861
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017L034"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             33.2%;  Score 727;  DB 7;  Length 861;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  777;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           25 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 84

Qy          126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           85 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 144

Qy          186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          145 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 204

Qy          246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          205 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 264

Qy          306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
                |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db          265 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 324

Qy          366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          325 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 384

Qy          426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          385 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 444

Qy          486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          445 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 504

Qy          546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          505 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 564

Qy          606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          565 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 624

Qy          666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          625 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 684

Qy          726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          685 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 744

Qy          786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          745 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCC 802

RESULT 199
CD517161
LOCUS       CD517161                 864 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14375039 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30395288 5', mRNA sequence.
ACCESSION   CD517161
VERSION     CD517161.1  GI:31448879
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 864)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
          Plate: NDAM475  row: c  column: 09
          High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1. .864
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30395288"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             33.2%;  Score 727;  DB 6;  Length 864;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  727;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        779 TGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAA 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAA 60

Qy        839 GATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCA 120

Qy        899 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCT 180

Qy        959 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 240

Qy       1019 CACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 300

Qy       1079 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGG 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGG 360

Qy       1139 TCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCT 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCT 420

Qy       1199 CAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGT 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGT 480

Qy       1259 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 540

Qy       1319 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 600

Qy       1379 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 1438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGAT 660

Qy       1439 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGA 720

Qy       1499 CACCATT 1505
              |||||||
Db        721 CACCATT 727


RESULT 200
CR999018
LOCUS       CR999018                741 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR999018 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016J1615 5',
            mRNA sequence.
ACCESSION   CR999018
VERSION     CR999018.1  GI:68292903
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016J1615.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016J1615
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016J1615"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. .
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             33.1%;  Score 726;  DB 7;  Length 741;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  726;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1467 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 1526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATC 60

Qy       1527 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 1586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTG 120

Qy       1587 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 1646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 GAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTG 180

Qy       1647 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 1706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACT 240

Qy       1707 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 300

Qy       1767 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 360

Qy       1827 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 1886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGA 420

Qy       1887 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 1946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAG 480

Qy       1947 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 2006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCT 540

Qy       2007 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 2066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 600

Qy       2067 TACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAA 2126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TACAACATCCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAA 660

Qy       2127 GTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGC 2186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCCCCTTTTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGC 720

Qy       2187 ATTCCC 2192
              ||||||
Db        721 ATTCCC 726
```

RESULT 201
CR983470
LOCUS CR983470 773 bp mRNA linear EST 24-JUN-2005
DEFINITION CR983470 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017L2324 5', mRNA sequence.
ACCESSION CR983470
VERSION CR983470.1 GI:68221804
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 773)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9017L2324.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017L2324
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .773
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9017L2324"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9017"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"
ORIGIN

Query Match 33.1%; Score 726; DB 7; Length 773;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 60

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 120

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 180

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 240

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 300

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 360

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 420

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTA 480

Qy        735 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACC 540

Qy        795 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAA 600

Qy        855 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGC 660
```

```
Qy        915 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGA 720

Qy        975 GGTGGG 980
              ||||||
Db        721 GGTGGG 726


RESULT 202
BM557281
LOCUS       BM557281               1088 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6561810 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5547785
            5', mRNA sequence.
ACCESSION   BM557281
VERSION     BM557281.1  GI:18799150
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1088)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12255  row: k  column: 18
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1. .1088
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5547785"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.1%;  Score 726;  DB 3;  Length 1088;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  726;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        298 TGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTG 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 TGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTG 81

Qy        358 ACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGC 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 ACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGC 141

Qy        418 AGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACA 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 AGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACA 201

Qy        478 TCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 TCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCG 261

Qy        538 TGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGA 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 TGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGA 321

Qy        598 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCT 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCT 381

Qy        658 TTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGT 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 TTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGT 441

Qy        718 CCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 CCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 501

Qy        778 CTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 CTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAA 561

Qy        838 AGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 AGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 621

Qy        898 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 681

Qy        958 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 741

Qy       1018 TCACCC 1023
              ||||||
Db        742 TCACCC 747


RESULT 203
CD514337
LOCUS       CD514337               835 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14373367 NIH_MGC_181 Homo sapiens cDNA clone
            IMAGE:30409051 5', mRNA sequence.
ACCESSION   CD514337
VERSION     CD514337.1  GI:31446055
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 835)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM492  row: p  column: 20
            High quality sequence stop: 653.
FEATURES             Location/Qualifiers
     source          1. .835
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
                     /clone="IMAGE:30409051"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             33.1%;  Score 725;  DB 6;  Length 835;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  725;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        920 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 96

Qy        980 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 156

Qy       1040 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 216

Qy       1100 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 276

Qy       1160 CATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 336

Qy       1220 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGA 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGA 396

Qy       1280 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 456

Qy       1340 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 516

Qy       1400 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 576

Qy       1460 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 636

Qy       1520 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 696

Qy       1580 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 756

Qy       1640 CCTGG 1644
              |||||
Db        757 CCTGG 761


RESULT 204
BU172676
LOCUS       BU172676                 904 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7959608 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6165758
            5', mRNA sequence.
ACCESSION   BU172676
VERSION     BU172676.1  GI:22686660
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13524  row: p  column: 15
            High quality sequence stop: 676.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6165758"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.1%;  Score 725;  DB 5;  Length 904;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  725;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         71 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 60

Qy        131 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 120

Qy        191 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 180

Qy        251 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 240

Qy        311 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 300

Qy        371 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 360

Qy        431 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 420

Qy        491 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 480

Qy        551 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 540

Qy        611 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 600

Qy        671 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 660

Qy        731 GGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCAT 720

Qy        791 CACCC 795
              |||||
Db        721 CACCC 725
```

RESULT 205
CX164590
LOCUS CX164590 724 bp mRNA linear EST 23-DEC-2004
DEFINITION HESC2_18_E02.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone IMAGE:7468110 5', mRNA sequence.
ACCESSION CX164590
VERSION CX164590.1 GI:56794670
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 724)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM15766 row: i column: 04
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 724.

FEATURES Location/Qualifiers
source 1. .724
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7468110"
/sex="male"
/tissue_type="human embryonic stem cells differentiated to an early endodermal cell type"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_258"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early endodermal cell type. Cell line id and NIH Registry designation is BG01. Positive for GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP expression. Passage number 40. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

```
  Query Match             33.0%;  Score 724;  DB 8;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 180

Qy        200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 240

Qy        260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 300

Qy        320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 360

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 420

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 480

Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 540

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 600

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 660

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 720

Qy        740 TCTC 743
              ||||
Db        721 TCTC 724
```

RESULT 206
CX866221

```
LOCUS       CX866221                724 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_9_B01.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7471803 5', mRNA sequence.
ACCESSION   CX866221
VERSION     CX866221.1  GI:58547739
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15776  row: c  column: 01
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 724.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7471803"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             33.0%;  Score 724;  DB 8;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 180

Qy        200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 240

Qy        260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 300

Qy        320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 360

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 420

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 480

Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 540

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 600

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 660

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 720

Qy        740 TCTC 743
              ||||
Db        721 TCTC 724


RESULT 207
CX873220
LOCUS       CX873220                724 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_78_B11.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7487952 5', mRNA sequence.
ACCESSION   CX873220
VERSION     CX873220.1  GI:58556394
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
           Plate: LLAM15818  row: c  column: 22
           Seq primer: JENREV (CAGGAAACAGCTATGACC)
           High quality sequence stop: 724.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7487952"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             33.0%;  Score 724;  DB 8;  Length 724;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 180

Qy        200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 240

Qy        260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 300

Qy        320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 360

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 420

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 480

Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 540

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        541 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 600

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 660

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 720

Qy        740 TCTC 743
              ||||
Db        721 TCTC 724


RESULT 208
CN415660
LOCUS       CN415660                742 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000424179406 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN415660
VERSION     CN415660.1  GI:47403254
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 742)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 742   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .742
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN

  Query Match             33.0%;  Score 724;  DB 7;  Length 742;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 GATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCT 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 GATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCT 78

Qy         79 TCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCG 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 TCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCG 138

Qy        139 AGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGA 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 AGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGA 198
```

```
Qy        199 TCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAG 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAG 258

Qy        259 AGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACAC 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 AGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACAC 318

Qy        319 TCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAG 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAG 378

Qy        379 CAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 CAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAA 438

Qy        439 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 AGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGC 498

Qy        499 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGA 558

Qy        559 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAG 618

Qy        619 ATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 ATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAG 678

Qy        679 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 738

Qy        739 GTCT 742
              ||||
Db        739 GTCT 742
```

```
RESULT 209
CR993103
LOCUS       CR993103                849 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR993103 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017G1819 5',
            mRNA sequence.
ACCESSION   CR993103
VERSION     CR993103.1  GI:68286988
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 849)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017G1819.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017G1819
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .849
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017G1819"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN
```

```
  Query Match             33.0%;  Score 724;  DB 7;  Length 849;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTC 60

Qy         81 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAG 120

Qy        141 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATC 180

Qy        201 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAG 240

Qy        261 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTC 300

Qy        321 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCA 360

Qy        381 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAG 420

Qy        441 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCAC 480

Qy        501 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACC 540

Qy        561 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGAT 600

Qy        621 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGAT 660

Qy        681 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 740
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGT 720

Qy        741 CTCA 744
                ||||
Db        721 CTCA 724


RESULT 210
BQ220564
LOCUS       BQ220564                875 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7595085 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062611
            5', mRNA sequence.
ACCESSION   BQ220564
VERSION     BQ220564.1  GI:20401964
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 875)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13334  row: f  column: 20
            High quality sequence stop: 562.
FEATURES             Location/Qualifiers
     source          1. .875
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6062611"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.0%;  Score 724;  DB 3;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         25 GGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGA 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGA 70

Qy         85 AGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATG 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 AGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATG 130

Qy        145 TCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTG 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 TCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTG 190

Qy        205 CTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 CTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 250

Qy        265 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 310

Qy        325 GCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGA 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 GCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGA 370

Qy        385 TCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 430

Qy        445 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 490

Qy        505 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 550

Qy        565 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 CCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 610

Qy        625 AAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 AAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 670

Qy        685 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 GTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 730

Qy        745 GAGG 748
                ||||
Db        731 GAGG 734


RESULT 211
BU153538
LOCUS       BU153538                883 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_7781992 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6138198
            5', mRNA sequence.
ACCESSION   BU153538
VERSION     BU153538.1  GI:22667070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13453  row: d  column: 07
            High quality sequence stop: 638.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6138198"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
```

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 33.0%; Score 724; DB 5; Length 883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy          76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy         136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy         196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy         376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy         436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy         496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy         556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy         616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy         676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy         736 TCCG 739
               ||||
Db         721 TCCG 724
```

RESULT 212
BQ434497
LOCUS BQ434497 900 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7892447 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6158601 5', mRNA sequence.
ACCESSION BQ434497
VERSION BQ434497.1 GI:21173573
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13506 row: f column: 10
High quality sequence stop: 612.
FEATURES Location/Qualifiers
source 1. .900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6158601"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
ORIGIN

Query Match 33.0%; Score 724; DB 5; Length 900;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy          76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy         136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy         196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy         256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy         316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy         376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy         436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy         496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy         556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 735
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTAC 720

Qy        736 TCCG 739
                ||||
Db        721 TCCG 724


RESULT 213
BU195402
LOCUS       BU195402                905 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7952229 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6171931
            5', mRNA sequence.
ACCESSION   BU195402
VERSION     BU195402.1  GI:22709386
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13541  row: a  column: 20
            High quality sequence stop: 600.
FEATURES             Location/Qualifiers
     source          1. .905
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6171931"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             33.0%;  Score 724;  DB 5;  Length 905;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  724;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 TAGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTC 85

Qy         61 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 ACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTG 145

Qy        121 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        146 AGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTC 205

Qy        181 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTG 265

Qy        241 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 ACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 325

Qy        301 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 AAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACA 385

Qy        361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 445

Qy        421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 505

Qy        481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 565

Qy        541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 625

Qy        601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 685

Qy        661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 745

Qy        721 CCTT 724
                ||||
Db        746 CCTT 749


RESULT 214
BQ670207
LOCUS       BQ670207                927 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AGENCOURT_8062804 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6211463
            5', mRNA sequence.
ACCESSION   BQ670207
VERSION     BQ670207.1  GI:21781041
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 927)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2373  row: p  column: 24
            High quality sequence stop: 659.
FEATURES             Location/Qualifiers
     source          1. .927
```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6211463"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_102"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 33.0%; Score 724; DB 5; Length 927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        548 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 60

Qy        608 AAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAA 120

Qy        668 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 180

Qy        728 CCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 240

Qy        788 CATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 300

Qy        848 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 360

Qy        908 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 420

Qy        968 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 480

Qy       1028 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 540

Qy       1088 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 1147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 600

Qy       1148 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 1207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 660

Qy       1208 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAG 1267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAG 720

Qy       1268 TGAC 1271
              ||||
Db        721 TGAC 724
```

RESULT 215
BX414032/c
LOCUS BX414032 889 bp mRNA linear EST 03-MAY-2004
DEFINITION BX414032 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA clone CS0DG006YB15 3-PRIME, mRNA sequence.
ACCESSION BX414032
VERSION BX414032.2 GI:46953032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30763402.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 6328.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAK046AB09NM1&c=6328.r.
FEATURES Location/Qualifiers
source 1. .889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG006YB15"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 33.0%; Score 723; DB 5; Length 889;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1415 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         773 GGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCAT 714

Qy       1475 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         713 CACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAA 654

Qy       1535 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 1594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         653 GGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGG 594

Qy       1595 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         593 ACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCT 534

Qy       1655 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 1714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         533 TAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGT 474

Qy       1715 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        473 GGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCC 414

Qy       1775 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 TCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTC 354

Qy       1835 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 TGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGAT 294

Qy       1895 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGA 234

Qy       1955 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 2014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCA 174

Qy       2015 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 2074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACAT 114

Qy       2075 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 2134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CCAGAAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCC 54

Qy       2135 TTT 2137
              |||
Db         53 TTT 51
```

```
RESULT 216
BU151351
LOCUS       BU151351                 860 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8726517 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6341111
            5', mRNA sequence.
ACCESSION   BU151351
VERSION     BU151351.1  GI:22664883
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 860)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2538  row: j  column: 24
            High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .860
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6341111"
                     /tissue_type="neuroblastoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_47"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match             32.9%;  Score 722;  DB 5;  Length 860;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         17 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 76
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGAT 63

Qy         77 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCAT 123

Qy        137 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCT 183

Qy        197 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAA 243

Qy        257 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 AGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGAC 303

Qy        317 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAA 363

Qy        377 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGG 423

Qy        437 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 AAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCT 483

Qy        497 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAA 543

Qy        557 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCA 603

Qy        617 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 AGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGA 663

Qy        677 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACT 723

Qy        737 CC 738
              ||
Db        724 CC 725
```

```
RESULT 217
CT000659
LOCUS       CT000659                 789 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT000659 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017J0312 5',
            mRNA sequence.
ACCESSION   CT000659
VERSION     CT000659.1  GI:68294542
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 789)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017J0312.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017J0312
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .789
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017J0312"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             32.9%;  Score 721;  DB 7;  Length 789;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  721;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         24 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 83
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTG 60

Qy         84 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAAT 120

Qy        144 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTT 180

Qy        204 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 240

Qy        264 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 300

Qy        324 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAG 360

Qy        384 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 420

Qy        444 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 480

Qy        504 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 540

Qy        564 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAG 600

Qy        624 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGT 660

Qy        684 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 720

Qy        744 A 744
              |
Db        721 A 721


RESULT 218
DR422128
LOCUS       DR422128                736 bp    mRNA    linear   EST 29-JUN-2005
DEFINITION  nav08d04.y1 Human pterygium. Unnormalized (nav) Homo sapiens cDNA
            clone nav08d04 5', mRNA sequence.
ACCESSION   DR422128
VERSION     DR422128.1  GI:68324144
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   Wistow,G., Jaworski,C., Aryankalayil-John,M., Rowsey,J.J., Cox,C.,
            Reid,T., Dushku,N. and Carper, D.
  TITLE     NEIBank analysis of Human pterygium
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 08  row: d  column: 04
            Seq primer: Universal M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .736
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="nav08d04"
                     /tissue_type="Pterygium"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human pterygium. Unnormalized (nav)"
                     /note="Organ: Eye; Vector: pCMVSport6; RNA was extracted
                     from 9 pooled human pterygia.  A directionally cloned cDNA
                     library in the pCMVSPORT6 vector (Invitrogen) was
                     constructed at Bioserve Biotechnology (Laurel MD)
                     essentially following the protocols of the SuperScript
                     Plasmid System, full details of which are contained in the
```

manufacturer's Instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adapter
[5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed at
the NIH Intramural Sequencing Center (NISC). Analyzed data
available through http://neibank.nei.nih.gov."
ORIGIN

Query Match 32.8%; Score 720; DB 8; Length 736;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1164 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 1223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 CAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 76

Qy       1224 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 1283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 AAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAAT 136

Qy       1284 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 1343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTT 196

Qy       1344 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 GCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 256

Qy       1404 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 1463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 ACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACT 316

Qy       1464 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 GGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAG 376

Qy       1524 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 ATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAG 436

Qy       1584 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 1643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 CTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTG 496

Qy       1644 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 1703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 GTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATC 556

Qy       1704 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 ACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 616

Qy       1764 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 1823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 GAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGA 676

Qy       1824 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 CGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTC 736
```

RESULT 219
BQ961879
LOCUS BQ961879 813 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8817704 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6379455
5', mRNA sequence.
ACCESSION BQ961879
VERSION BQ961879.1 GI:22377357
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 813)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2566 row: h column: 16
High quality sequence stop: 694.
FEATURES Location/Qualifiers
source 1. .813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6379455"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
ORIGIN

Query Match 32.8%; Score 720; DB 5; Length 813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        161 AGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGA 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGA 60

Qy        221 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCT 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCT 120

Qy        281 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 180

Qy        341 TGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGG 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGG 240

Qy        401 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCAC 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCAC 300

Qy        461 CCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGG 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGG 360

Qy        521 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCC 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCC 420

Qy        581 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGA 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGA 480
```

```
Qy        641 TCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTA 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTA 540

Qy        701 CAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAAT 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAAT 600

Qy        761 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTAT 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTAT 660

Qy        821 CGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTT 720
```

RESULT 220
BQ935362
LOCUS BQ935362 910 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8794222 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6370488 5', mRNA sequence.
ACCESSION BQ935362
VERSION BQ935362.1 GI:22350745
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2543 row: c column: 01
High quality sequence stop: 659.

FEATURES Location/Qualifiers
source 1. .910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6370488"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 32.8%; Score 720; DB 5; Length 910;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        788 CATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 CATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 79

Qy        848 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 139

Qy        908 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 199

Qy        968 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 259

Qy       1028 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 319

Qy       1088 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 1147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 379

Qy       1148 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 1207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 439

Qy       1208 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAG 1267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAG 499

Qy       1268 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCA 1327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCA 559

Qy       1328 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 619

Qy       1388 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 1447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTT 679

Qy       1448 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 1507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGA 739
```

RESULT 221
CX165664
LOCUS CX165664 719 bp mRNA linear EST 23-DEC-2004
DEFINITION HESC2_31_C06.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone IMAGE:7469197 5', mRNA sequence.
ACCESSION CX165664
VERSION CX165664.1 GI:56795744
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 719)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

```
           http://image.llnl.gov
           Plate: LLAM15769  row: f  column: 11
           Seq primer: JENREV (CAGGAAACAGCTATGACC)
           High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7469197"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
                     Genomics (Frederick, MD). Sequence ends have been trimmed
                     to exclude vector and regions below Phred quality 16.
                     Three-prime sequences are presented as their reverse
                     complement and have been trimmed to exclude polyA. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             32.8%;  Score 719;  DB 8;  Length 719;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  719;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTT 60

Qy         80 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGA 120

Qy        140 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGAT 180

Qy        200 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGA 240

Qy        260 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACT 300

Qy        320 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGC 360

Qy        380 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 420

Qy        440 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 480

Qy        500 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 540

Qy        560 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGA 600

Qy        620 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 660

Qy        680 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 719


RESULT 222
DR155702
LOCUS       DR155702                719 bp    mRNA    linear   EST 16-JUN-2005
DEFINITION  HESC2_57_C09.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7964251 5', mRNA sequence.
ACCESSION   DR155702
VERSION     DR155702.1  GI:67867862
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM17069  row: e  column: 17
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7964251"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
                     Genomics (Frederick, MD). Sequence ends have been trimmed
```

to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 32.8%; Score 719; DB 8; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        817 CTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGA 60

Qy        877 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 120

Qy        937 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 180

Qy        997 AGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATG 240

Qy       1057 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTG 300

Qy       1117 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 360

Qy       1177 CCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTG 420

Qy       1237 GTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGA 480

Qy       1297 TCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGC 540

Qy       1357 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGG 600

Qy       1417 TGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 660

Qy       1477 CTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAG 719
```

RESULT 223
DR759813
LOCUS DR759813 769 bp mRNA linear EST 21-JUL-2005
DEFINITION HESC4_108_C01.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone IMAGE:7968868 5', mRNA sequence.
ACCESSION DR759813
VERSION DR759813.1 GI:71052513
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM 17081 row: f column: 02
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 769.
FEATURES Location/Qualifiers
source 1. .769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7968868"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 32.8%; Score 718; DB 8; Length 769;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 768; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTACTCCGTCTCAGAG 720

Qy        748 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCT 769
```

```
RESULT 224
BU164691
LOCUS       BU164691                865 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7782181 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6138138
            5', mRNA sequence.
ACCESSION   BU164691
VERSION     BU164691.1  GI:22678643
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 865)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13453  row: a  column: 19
            High quality sequence stop: 548.
FEATURES             Location/Qualifiers
     source          1. .865
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6138138"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN
```

```
  Query Match           32.8%;  Score 718;  DB 5;  Length 865;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  718;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 60

Qy         89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 120

Qy        149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 180

Qy        209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 240

Qy        269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 300

Qy        329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 360

Qy        389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 420

Qy        449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 480

Qy        509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 540

Qy        569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 600

Qy        629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 660

Qy        689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 718
```

```
RESULT 225
BU171951
LOCUS       BU171951                884 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7932606 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6157287
            5', mRNA sequence.
ACCESSION   BU171951
VERSION     BU171951.1  GI:22685935
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13502  row: o  column: 16
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .884
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6157287"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match            32.8%;  Score 718;  DB 5;  Length 884;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  718;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 60

Qy         91 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 120

Qy        151 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 180

Qy        211 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 240

Qy        271 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 300

Qy        331 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 360

Qy        391 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 420

Qy        451 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 480

Qy        511 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 540

Qy        571 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 600

Qy        631 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 660

Qy        691 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGG 718


RESULT 226
CD251905
LOCUS       CD251905                894 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14204071 NIH_MGC_180 Homo sapiens cDNA clone
            IMAGE:30383213 5', mRNA sequence.
ACCESSION   CD251905
VERSION     CD251905.1  GI:31012371
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM446  row: l  column: 06
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30383213"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_180"
                     /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI;
                     Site_2: EcoRV (destroyed); Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.68 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match            32.8%;  Score 718;  DB 6;  Length 894;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  718;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         33 TTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCG 92

Qy         82 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 TGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGA 152

Qy        142 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 ATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCT 212

Qy        202 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGT 272

Qy        262 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        273 CCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCA 332

Qy        322 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 CTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAA 392

Qy        382 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 AGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGC 452

Qy        442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 512

Qy        502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 572

Qy        562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 632

Qy        622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 692

Qy        682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCG 750


RESULT 227
BX442070
LOCUS       BX442070                1008 bp    mRNA    linear   EST 04-MAY-2004
DEFINITION  BX442070 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF025YI16 5-PRIME, mRNA sequence.
ACCESSION   BX442070
VERSION     BX442070.2  GI:47009028
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1008)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30781989.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DF025BE08QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1008
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF025YI16"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match           32.8%;  Score 718;  DB 5;  Length 1008;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  768;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        127 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACC 60

Qy        187 AGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACA 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACA 120

Qy        247 ACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCT 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCT 180

Qy        307 TCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCG 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCG 240

Qy        367 AGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGA 426
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        241 AGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGAYCAGCAGAGGTTGA 300

Qy        427 TCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAG 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAG 360

Qy        487 AGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCC 420

Qy        547 TGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGG 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGG 480

Qy        607 CAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAA 540

Qy        667 AACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGC 600

Qy        727 ACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 660

Qy        787 CCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAG 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAG 720

Qy        847 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 769


RESULT 228
CN415659
LOCUS       CN415659                717 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000582991437 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN415659
VERSION     CN415659.1  GI:47403253
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 717)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 717   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .717
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match           32.7%;  Score 717;  DB 7;  Length 717;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  717;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 744
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 717


RESULT 229
CN415629
LOCUS       CN415629                725 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000599934083 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN415629
VERSION     CN415629.1  GI:47403223
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
   PUBMED   15146197
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 725   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match           32.7%;  Score 717;  DB 7;  Length 725;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  717;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 68

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 128

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 188

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 248

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 308
```

```
Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 368

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 428

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 488

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 548

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 608

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 668

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 725
```

```
RESULT 230
CD556735
LOCUS       CD556735                897 bp    mRNA    linear   EST 11-JUN-2003
DEFINITION  AGENCOURT_14380657 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30393477 5', mRNA sequence.
ACCESSION   CD556735
VERSION     CD556735.1  GI:31582803
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM470  row: g  column: 22
            High quality sequence stop: 623.
FEATURES             Location/Qualifiers
     source          1. .897
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30393477"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN
```

```
  Query Match           32.7%;  Score 717;  DB 6;  Length 897;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  717;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        594 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTG 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTG 60

Qy        654 ATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAA 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAA 120

Qy        714 GAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACA 180

Qy        774 CTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAA 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAA 240

Qy        834 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGA 300

Qy        894 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTG 360

Qy        954 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAG 420

Qy       1014 ACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAA 480

Qy       1074 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAA 540

Qy       1134 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTC 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTC 600

Qy       1194 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTC 660

Qy       1254 GAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAGGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 717
```

```
RESULT 231
BU175009
LOCUS       BU175009                970 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_8073968 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6087415
            5', mRNA sequence.
ACCESSION   BU175009
VERSION     BU175009.1  GI:22688993
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2323  row: p  column: 08
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1. .970
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6087415"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_102"
                     /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California, Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life
                     Technologies). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           32.7%;  Score 717;  DB 5;  Length 970;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  717;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        548 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGC 60

Qy        608 AAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAA 120

Qy        668 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCA 180

Qy        728 CCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 240

Qy        788 CATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGA 300

Qy        848 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 360

Qy        908 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 420

Qy        968 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 480

Qy       1028 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 540

Qy       1088 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 1147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 600

Qy       1148 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 1207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGG 660

Qy       1208 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCC 717


RESULT 232
CF131936
LOCUS       CF131936                716 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-aws-m-13-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30553932 5', mRNA sequence.
ACCESSION   CF131936
VERSION     CF131936.1  GI:33214693
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  PUBMED    8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30553932"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with Not I and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

  Query Match           32.7%;  Score 716;  DB 6;  Length 716;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  716;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 716


RESULT 233
CT004302
LOCUS       CT004302                 716 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT004302 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016A034 5',
            mRNA sequence.
ACCESSION   CT004302
VERSION     CT004302.1  GI:68298185
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016A034.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016A034
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9016A034"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9016"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             32.7%;  Score 716;  DB 7;  Length 716;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  716;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600
```

```
Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCT 716


RESULT 234
CX872917
LOCUS       CX872917                716 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_74_F09.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7487756 5', mRNA sequence.
ACCESSION   CX872917
VERSION     CX872917.1  GI:58556091
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 716)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15817  row: k  column: 18
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 716.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7487756"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match           32.7%;  Score 716;  DB 8;  Length 716;
  Best Local Similarity 100.0%;  Pred. No. 0;
  Matches  716;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGA 60

Qy         88 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCA 120

Qy        148 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTG 180

Qy        208 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCC 240

Qy        268 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCA 300

Qy        328 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCC 360

Qy        388 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGG 420

Qy        448 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGC 480

Qy        508 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCC 540

Qy        568 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 600

Qy        628 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTA 660

Qy        688 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 716


RESULT 235
BU193795
LOCUS       BU193795                816 bp    mRNA    linear   EST 04-SEP-2002
DEFINITION  AGENCOURT_7945646 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6149985
            5', mRNA sequence.
ACCESSION   BU193795
VERSION     BU193795.1  GI:22707779
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 816)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13483  row: o  column: 10
           High quality sequence stop: 548.
FEATURES             Location/Qualifiers
     source          1. .816
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6149985"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             32.7%;  Score 716;  DB 5;  Length 816;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  716;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGAC 60

Qy         89 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAA 120

Qy        149 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGG 180

Qy        209 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCT 240

Qy        269 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAA 300

Qy        329 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCA 360

Qy        389 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGA 420

Qy        449 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCT 480

Qy        509 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCT 540

Qy        569 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGG 600

Qy        629 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTAC 660

Qy        689 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCA 716
```

```
RESULT 236
BU150027
LOCUS       BU150027                888 bp    mRNA    linear   EST 03-SEP-2002
DEFINITION  AGENCOURT_8228879 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6183554 5', mRNA sequence.
ACCESSION   BU150027
VERSION     BU150027.1  GI:22663570
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13571  row: f  column: 03
           High quality sequence stop: 603.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6183554"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match             32.7%;  Score 716;  DB 5;  Length 888;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  716;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         31 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTC 60

Qy         91 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGG 120

Qy        151 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAA 180

Qy        211 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGC 240

Qy        271 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 330
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGA 300

Qy        331 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGG 360

Qy        391 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAG 420

Qy        451 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCC 480

Qy        511 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCG 540

Qy        571 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCA 600

Qy        631 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCC 660

Qy        691 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGA 716


RESULT 237
CT003348
LOCUS       CT003348                 920 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT003348 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017B1439 5',
            mRNA sequence.
ACCESSION   CT003348
VERSION     CT003348.1  GI:68297231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 920)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017B1439.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017B1439
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1..920
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017B1439"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             32.7%;  Score 716;  DB 7;  Length 920;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  766;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 96

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 156

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 216

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 276

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        277 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 336

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 396

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 456

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 516

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 576

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 636

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 696

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 756

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 832
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        757 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 803


RESULT 238
```

```
CR988580
LOCUS       CR988580                930 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CR988580 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017A1513 5',
            mRNA sequence.
ACCESSION   CR988580
VERSION     CR988580.1  GI:68282465
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017A1513.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017A1513
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017A1513"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match           32.7%;  Score 716;  DB 7;  Length 930;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  766;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 111

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 171

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 231

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 291

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        292 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 351

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 411

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 471

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 531

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 591

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 651

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 771

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 832
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 818


RESULT 239
CR977118
LOCUS       CR977118                936 bp    mRNA    linear   EST 24-JUN-2005
DEFINITION  CR977118 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016A061 5',
            mRNA sequence.
ACCESSION   CR977118
VERSION     CR977118.1  GI:68215435
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 936)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016A061.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016A061
            contact RZPD (product- support@rzpd.de) for further information.
```

Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.

FEATURES Location/Qualifiers
source 1. .936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016A061"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 32.7%; Score 716; DB 7; Length 936;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 766; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 111

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 171

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 231

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 291

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        292 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 351

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 411

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 471

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 531

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 591

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 651

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 771

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 832
               ||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAA 818
```

RESULT 240
BQ889973
LOCUS BQ889973 949 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_7982318 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone IMAGE:6186134 5', mRNA sequence.
ACCESSION BQ889973
VERSION BQ889973.1 GI:22281987
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 949)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM13578 row: a column: 15
High quality sequence stop: 573.

FEATURES Location/Qualifiers
source 1. .949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6186134"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 32.7%; Score 716; DB 5; Length 949;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 766; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 75

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 135

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 195
```

```
Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 255

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 315

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        316 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 375

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 435

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 495

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 555

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 615

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 675

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        676 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 735

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTG 798
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        736 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTG 782


RESULT 241
BQ049067
LOCUS       BQ049067                1006 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6794970 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5786835
            5', mRNA sequence.
ACCESSION   BQ049067
VERSION     BQ049067.1  GI:19808407
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1006)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12876  row: d  column: 04
            High quality sequence stop: 652.
FEATURES             Location/Qualifiers
     source          1. .1006
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5786835"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             32.7%;  Score 716;  DB 3;  Length 1006;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  766;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 63

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 123

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 183

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 243

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 303

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 363

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 423

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 483

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 543

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 603

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 663

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        664 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAAGTGG 723

Qy        752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTG 798
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTG 770


RESULT 242
BM477326
LOCUS       BM477326                1093 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6484868 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5554104
```

```
            5', mRNA sequence.
ACCESSION   BM477326
VERSION     BM477326.1  GI:18526368
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1093)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12272  row: c  column: 01
            High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1. .1093
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5554104"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            32.7%;  Score 716;  DB 3;  Length 1093;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  716;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         17 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 76

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 136

Qy        152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 196

Qy        212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 256

Qy        272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 316

Qy        332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 376

Qy        392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 436

Qy        452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 496

Qy        512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 556

Qy        572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        557 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 616

Qy        632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 676

Qy        692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAG 732


RESULT 243
BM476508
LOCUS       BM476508                1070 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6476298 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5553443
            5', mRNA sequence.
ACCESSION   BM476508
VERSION     BM476508.1  GI:18525550
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1070)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12270  row: g  column: 12
            High quality sequence stop: 599.
FEATURES             Location/Qualifiers
     source          1. .1070
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5553443"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            32.6%;  Score 715;  DB 3;  Length 1070;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  765;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         32 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 GTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCT 67

Qy         92 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 151
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          68 GACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGC 127

Qy         152 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 211
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         128 AAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAA 187

Qy         212 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 271
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         188 ACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCA 247

Qy         272 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 331
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         248 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGAC 307

Qy         332 CATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 391
                |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         308 CATCACCCTTGAGGTCGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGA 367

Qy         392 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 451
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         368 CAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGA 427

Qy         452 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 511
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         428 TGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCG 487

Qy         512 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 571
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         488 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGA 547

Qy         572 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 631
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         548 GGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCAT 607

Qy         632 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 691
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         608 TCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCT 667

Qy         692 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 751
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         668 GTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGG 727

Qy         752 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 797
                |||||||||||||||||||||||||||||||||||||||||||||
Db         728 GATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTT 773


RESULT 244
AL516880
LOCUS       AL516880               1102 bp    mRNA    linear   EST 23-MAR-2004
DEFINITION  AL516880 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            CS0DA007YA09 5-PRIME, mRNA sequence.
ACCESSION   AL516880
VERSION     AL516880.3  GI:45653489
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 1102)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 13, 2001 this sequence version replaced gi:30491945.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6328.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DA007AA05QP1&c=6328.r.
FEATURES             Location/Qualifiers
     source          1. .1102
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA007YA09"
                     /tissue_type="NEUROBLASTOMA"
                     /clone_lib="Homo sapiens NEUROBLASTOMA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             32.6%;  Score 715;  DB 1;  Length 1102;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  815;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          49 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          38 GCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCA 97

Qy         109 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          98 CCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGG 157

Qy         169 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         158 AAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGC 217

Qy         229 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         218 GCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCA 277

Qy         289 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGG 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         278 GAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCG 337

Qy         349 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         338 AGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTC 397

Qy         409 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         398 CTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTG 457

Qy         469 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         458 ACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGC 517

Qy         529 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 588
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         518 AGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACA 577

Qy         589 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 648
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         578 CCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGA 637

Qy         649 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 708
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         638 GGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 697

Qy         709 AGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 768
               |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         698 AGAAAGAGTCCACCYTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGA 757

Qy         769 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 828
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 AGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACG 817

Qy        829 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCC 865
              |||||||||||||||||||||||||||||||||||||
Db        818 TCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCC 854


RESULT 245
CX873027
LOCUS       CX873027                 714 bp    mRNA    linear   EST 03-FEB-2005
DEFINITION  HESC4_75_H11.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
            IMAGE:7487879 5', mRNA sequence.
ACCESSION   CX873027
VERSION     CX873027.1  GI:58556201
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 714)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15817  row: p  column: 21
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 714.
FEATURES             Location/Qualifiers
     source          1..714
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7487879"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             32.6%;  Score 714;  DB 8;  Length 714;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  714;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1479 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 1538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAA 60

Qy       1539 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 1598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGC 120

Qy       1599 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGA 1658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGA 180

Qy       1659 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 1718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAG 240

Qy       1719 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 1778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCT 300

Qy       1779 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 1838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGAC 360

Qy       1839 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 1898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAG 420

Qy       1899 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 1958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACC 480

Qy       1959 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 2018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGG 540

Qy       2019 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 2078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAG 600

Qy       2079 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 2138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAGAGTCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTT 660

Qy       2139 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 714


RESULT 246
BQ934245
LOCUS       BQ934245                 906 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8878940 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6202151 5', mRNA sequence.
ACCESSION   BQ934245
VERSION     BQ934245.1  GI:22349628
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
```

```
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13619  row: 1  column: 24
          High quality sequence stop: 601.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6202151"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sciatic_nerve"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match             32.5%;  Score 713;  DB 5;  Length 906;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches  813;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        836 AAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAA 60

Qy        896 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCA 120

Qy        956 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGAC 180

Qy       1016 CATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 1075
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGA 240

Qy       1076 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGA 300

Qy       1136 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCG 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCG 360

Qy       1196 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGA 420

Qy       1256 GGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCAT 480

Qy       1316 CCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 1375
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCT 540

Qy       1376 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGG 600

Qy       1436 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAG 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAG 660

Qy       1496 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 720

Qy       1556 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 1615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAA 780

Qy       1616 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 1650
              ||||||||||||||||||||||||||||||||||
Db        781 CATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCC 815


RESULT 247
AU130085
LOCUS       AU130085                842 bp    mRNA    linear   EST 01-AUG-2002
DEFINITION  AU130085 NT2RP3 Homo sapiens cDNA clone NT2RP3000224 5', mRNA
            sequence.
ACCESSION   AU130085
VERSION     AU130085.1  GI:10990439
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 842)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
            Isogai,T.
  TITLE     HRI human cDNA project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .842
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3000224"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /clone_lib="NT2RP3"
                     /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
                     cells after 2-weeks retinoic acid (RA) induction"
ORIGIN

  Query Match             32.5%;  Score 712;  DB 1;  Length 842;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  712;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           2 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 61
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AGTTCCGTCGCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCA 60

Qy          62 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 121
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGA 120

Qy         122 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 181
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCC 180
```

```
Qy        182 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGA 240

Qy        242 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCA 300

Qy        302 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACAC 360

Qy        362 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAG 420

Qy        422 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCA 480

Qy        482 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAA 540

Qy        542 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGT 600

Qy        602 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGC 660

Qy        662 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAA 713
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAA 712
```

RESULT 248
BQ421393
LOCUS BQ421393 875 bp mRNA linear EST 23-MAY-2002
DEFINITION AGENCOURT_7771428 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6019361 5', mRNA sequence.
ACCESSION BQ421393
VERSION BQ421393.1 GI:21116708
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13221 row: 1 column: 18
High quality sequence stop: 540.

```
FEATURES             Location/Qualifiers
     source          1. .875
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6019361"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
ORIGIN

  Query Match            32.5%;  Score 712;  DB 5;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  712;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        453 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 73

Qy        513 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 133

Qy        573 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATT 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATT 193

Qy        633 CCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 253

Qy        693 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGG 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGG 313

Qy        753 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGT 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 ATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGT 373

Qy        813 GACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAG 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 GACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCATTCCTCCTGACCAG 433

Qy        873 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 CAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAAC 493

Qy        933 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 ATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 553

Qy        993 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAG 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAGTGACACCATTGAG 613

Qy       1053 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 AATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATC 673

Qy       1113 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 1164
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 TTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCC 725
```

RESULT 249
CD515652
LOCUS CD515652 875 bp mRNA linear EST 06-JUN-2003
DEFINITION AGENCOURT_14364897 NIH_MGC_181 Homo sapiens cDNA clone IMAGE:30407992 5', mRNA sequence.
ACCESSION CD515652
VERSION CD515652.1 GI:31447370
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 875)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM490  row: d  column: 17
            High quality sequence stop: 711.
FEATURES             Location/Qualifiers
     source          1. .875
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30407992"
                     /tissue_type="White Matter"
                     /dev_stage="Unknown"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_181"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
                     (destroyed); Library is oligo-dT primed and directionally
                     cloned (EcoRV site is destroyed upon cloning). Average
                     insert size 1.42 kb. Library was constructed by
                     (Invitrogen). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             32.5%;  Score 712;  DB 6;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  712;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        920 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 GTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGG 96

Qy        980 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGAG 156

Qy       1040 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 TGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCA 216

Qy       1100 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 GCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAA 276

Qy       1160 CATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTT 336

Qy       1220 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGA 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTGA 396

Qy       1280 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGAT 456

Qy       1340 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 CTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGA 516

Qy       1400 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATCTTCGTGAAGACCCT 576

Qy       1460 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 GACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGC 636

Qy       1520 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 AAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAA 696

Qy       1580 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 1631
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 ACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCC 748


RESULT 250
DR156196
LOCUS       DR156196                711 bp    mRNA    linear   EST 16-JUN-2005
DEFINITION  HESC2_63_B02.g1_A035 NIH_MGC_258 Homo sapiens cDNA clone
            IMAGE:7964597 5', mRNA sequence.
ACCESSION   DR156196
VERSION     DR156196.1  GI:67868356
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 711)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM 17070  row: d  column: 03
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 711.
FEATURES             Location/Qualifiers
     source          1. .711
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7964597"
                     /sex="male"
                     /tissue_type="human embryonic stem cells differentiated to
                     an early endodermal cell type"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_258"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early endodermal cell type. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     GATA4, MixL1, Msx1, HNF4alpha expression; negative for AFP
                     expression. Passage number 40. cDNA primed using oligo-dT
                     primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and
                     cloned into the EcoRV/NotI sites of pExpress-1. This
                     primary library is non-normalized (normalized primary
                     library is NIH_MGC_259). It was constructed by Express
```

Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 32.4%; Score 711; DB 8; Length 711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         15 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAG 60

Qy         75 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACC 120

Qy        135 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGG 180

Qy        195 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAG 240

Qy        255 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAG 300

Qy        315 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTC 360

Qy        375 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCC 420

Qy        435 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACC 480

Qy        495 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGT 540

Qy        555 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATC 600

Qy        615 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTG 660

Qy        675 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711
```

RESULT 251
DR760405
LOCUS DR760405 711 bp mRNA linear EST 21-JUL-2005
DEFINITION HESC4_115_G05.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone IMAGE:7969835 5', mRNA sequence.
ACCESSION DR760405
VERSION DR760405.1 GI:71053105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 711)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: BresaGen, Inc.
cDNA Library Preparation: Express Genomics, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Laboratory for Genomics and Bioinformatics, University of Georgia
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM 17083 row: n column: 09
Seq primer: JENREV (CAGGAAACAGCTATGACC)
High quality sequence stop: 711.

FEATURES Location/Qualifiers
source 1. .711
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7969835"
/sex="male"
/tissue_type="embryonic stem"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH_MGC_262"
/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV; RNA obtained from human embryonic stem cells isolated from the inner cell mass of blastocyst stage embryos and differentiated to an early neural progenitor cell type. Cell line id and NIH Registry designation is BG01. Positive for Nestin and Musashi expression. Passage number 18. cDNA primed using oligo-dT primer: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. This primary library is non-normalized (normalized primary library is NIH_MGC_259). It was constructed by Express Genomics (Frederick, MD). Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Note: this is a Mammalian Gene Collection library."

ORIGIN

Query Match 32.4%; Score 711; DB 8; Length 711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         33 TTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTG 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTG 60

Qy         93 ACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCA 120

Qy        153 AAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAA 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAA 180

Qy        213 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCAC 240

Qy        273 CTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACC 300

Qy        333 ATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGAC 392
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGAC 360

Qy        393 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 452
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGAT 420

Qy        453 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 512
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGT 480

Qy        513 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 572
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAG 540

Qy        573 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATT 632
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATT 600

Qy        633 CCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 692
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTG 660

Qy        693 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 743
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTCTC 711


RESULT 252
CX782198
LOCUS       CX782198                762 bp    mRNA    linear   EST 02-MAR-2005
DEFINITION  HESC3_12_F12.g1_A036 NIH_MGC_260 Homo sapiens cDNA clone
            IMAGE:7476650 5', mRNA sequence.
ACCESSION   CX782198
VERSION     CX782198.1  GI:58298988
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15788  row: 1  column: 24
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 762.
FEATURES             Location/Qualifiers
     source          1. .762
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7476650"
                     /sex="male"
                     /tissue_type="human embryonic stem cells"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_260"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells isolated from
                     the inner cell mass of blastocyst stage embryos. Cell line
                     id and NIH Registry designation is BG01. Positive for
                     SSEA3, SSEA4, Tra 1-60, Tra 1-81, CD9, Alk Phos, Oct4 and
                     Nanog expression; negative for SSEA1 expression. Passage
                     number 21. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_261). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match             32.4%;  Score 711;  DB 8;  Length 762;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  761;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        679 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 738
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCC 60

Qy        739 GTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTG 798
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTG 120

Qy        799 AGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCA 858
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAAGACAAGGAAGGCA 180

Qy        859 TTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 918
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCC 240

Qy        919 TGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 978
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTG 300

Qy        979 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGA 1038
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAAGTGGAGCCGA 360

Qy       1039 GTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACC 1098
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACC 420

Qy       1099 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACA 1158
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACA 480

Qy       1159 ACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 1218
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACATCCAGAAAGAGTCCACCTTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCT 540

Qy       1219 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTG 1278
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAGGTGGAGCCGAGTGACACCATTG 600

Qy       1279 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGA 1338
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGA 660

Qy       1339 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAG 1398
                |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        661 TCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACNACATCCAGAAAG 720

Qy       1399 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 1440
                |||||||||||||||||||||||||||||||||||||||||
Db        721 AGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGC 762
```

RESULT 253
BQ879740
LOCUS BQ879740 942 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8122185 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone IMAGE:6179435 5', mRNA sequence.
ACCESSION BQ879740
VERSION BQ879740.1 GI:22271748
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 942)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13560 row: j column: 12
High quality sequence stop: 675.
FEATURES Location/Qualifiers
source 1. .942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179435"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
ORIGIN

Query Match 32.4%; Score 710; DB 5; Length 942;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        300 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGAC 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGAC 60

Qy        360 ACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAG 120

Qy        420 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATC 180

Qy        480 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTG 240

Qy        540 AAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAAT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAAT 300

Qy        600 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTT 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTT 360

Qy        660 GCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCC 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCC 420

Qy        720 ACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACT 480

Qy        780 GGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAG 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAG 540

Qy        840 ATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAG 600

Qy        900 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTG 660

Qy        960 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGG 710
```

RESULT 254
BQ881494
LOCUS BQ881494 851 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8137962 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone IMAGE:6183403 5', mRNA sequence.
ACCESSION BQ881494
VERSION BQ881494.1 GI:22273502
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13570 row: o column: 20
High quality sequence stop: 638.
FEATURES Location/Qualifiers
source 1. .851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6183403"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors:

5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 32.3%; Score 709; DB 5; Length 851;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         16 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAATGCAGA 60

Qy         76 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGACACCA 120

Qy        136 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCAGAGGC 180

Qy        196 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGA 240

Qy        256 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGA 300

Qy        316 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAACGTCA 360

Qy        376 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCG 420

Qy        436 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCC 480

Qy        496 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTA 540

Qy        556 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCC 600

Qy        616 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGG 660

Qy        676 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 724
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTT 709
```

RESULT 255
BM975559/c
LOCUS BM975559 726 bp mRNA linear EST 21-FEB-2003
DEFINITION UI-CF-EN1-acw-k-17-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone UI-CF-EN1-acw-k-17-0-UI 3', mRNA sequence.
ACCESSION BM975559
VERSION BM975559.1 GI:19593150
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .726
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-acw-k-17-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h
TAG_LIB=UI-CF-EN1
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 32.3%; Score 708; DB 3; Length 726;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1485 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATC 667

Qy       1545 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CCTCCTGACCAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTG 607

Qy       1605 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGG 547

Qy       1665 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        546 ATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGT 487

Qy       1725 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 GACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAG 427

Qy       1785 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAAC 367

Qy       1845 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 1904
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTC 307

Qy       1905 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAG 1964
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAG 247

Qy       1965 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATC 2024
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATC 187

Qy       2025 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 2084
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAG 127

Qy       2085 TCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTT 2144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCACTCTGCACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTT 67

Qy       2145 TCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 19
```

RESULT 256
CF132206
LOCUS CF132206 769 bp mRNA linear EST 05-AUG-2003
DEFINITION UI-HF-FQ0-awr-b-16-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone IMAGE:30555207 5', mRNA sequence.
ACCESSION CF132206
VERSION CF132206.1 GI:33215237
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30555207"
/tissue_type="Chondrosarcoma Lung Metastasis cell lines"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_215"
/note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

Query Match 32.3%; Score 708; DB 6; Length 769;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          11 GCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAAT 70
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCAGCCGGGATTTGGGTCGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTGACAAT 60

Qy          71 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCCAGTGA 120

Qy         131 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGACCAGCA 180

Qy         191 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 250
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACAT 240

Qy         251 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 310
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATCTTCGT 300

Qy         311 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 370
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATCGAGAA 360

Qy         371 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 430
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTT 420

Qy         431 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 490
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTC 480

Qy         491 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 550
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGAC 540

Qy         551 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 610
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAA 600

Qy         611 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 670
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACA 660

Qy         671 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTC 718
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTC 708
```

RESULT 257
BU557860/c
LOCUS BU557860 915 bp mRNA linear EST 16-SEP-2002

```
DEFINITION  AGENCOURT_10212402 NIH_MGC_107 Homo sapiens cDNA clone
            IMAGE:6586064 5', mRNA sequence.
ACCESSION   BU557860
VERSION     BU557860.1  GI:22908156
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2799  row: a  column: 08
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6586064"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_107"
                     /note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             32.3%;  Score 708;  DB 5;  Length 915;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  758;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1434 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 1493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCG 734

Qy       1494 AGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 1553
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 AGTGACCCCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGAC 674

Qy       1554 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 1613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 CAGCAGAGGTTGATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTAC 614

Qy       1614 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 1673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTTAGAGGTGGGATGCAGATC 554

Qy       1674 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 1733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 TTCGTGAAGACCCTGACTGGTAAGACCATCACTCTCGAAGTGGAGCCGAGTGACACCATT 494

Qy       1734 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 1793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 GAGAATGTCAAGGCAAAGATCCAAGACAAGGAAGGCATCCCTCCTGACCAGCAGAGGTTG 434

Qy       1794 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 1853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 ATCTTTGCTGGGAAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAA 374

Qy       1854 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 1913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 314

Qy       1914 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 1973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 254

Qy       1974 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 2033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGATCAGCAGAGGTTGATCTTTGCTGGG 194

Qy       2034 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 2093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 AAACAGCTGGAAGATGGACGCACCCTGTCTGACTACAACATCCAGAAAGAGTCCACTCTG 134

Qy       2094 CACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAAT 2153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 CACTTGGTCCTGCGCTTGAGGGGGGGTGTCTAAGTTTCCCCTTTTAAGGTTTCAACAAAT 74

Qy       2154 TTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 2192
             |||||||||||||||||||||||||||||||||||||||
Db          73 TTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCC 35


RESULT 258
CF131782
LOCUS       CF131782                783 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-FQ0-awr-p-07-0-UI.r1 NIH_MGC_215 Homo sapiens cDNA clone
            IMAGE:30555534 5', mRNA sequence.
ACCESSION   CF131782
VERSION     CF131782.1  GI:33214384
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 783)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .783
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30555534"
                     /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_MGC_215"
                     /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
```

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 32.3%; Score 707; DB 6; Length 783;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        52 GTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCC 111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GTGATCGTCACTTGACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCC 60

Qy       112 TCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TCGAGGTTGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAG 120

Qy       172 GCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCA 231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCATCCCTCCTGACCAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCA 180

Qy       232 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 291
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CCCTGTCTGACTACAACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAG 240

Qy       292 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGC 351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGC 300

Qy       352 CCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTG 411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CCAGTGACACCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTG 360

Qy       412 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACT 420

Qy       472 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 ACAACATCCAGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGA 480

Qy       532 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCA 591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TCTTCGTGAAGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCA 540

Qy       592 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGT 651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGT 600

Qy       652 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TGATCTTTGCCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGA 660

Qy       712 AAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 758
             ||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AAGAGTCCACCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAA 707
```

RESULT 259
BQ689319
LOCUS BQ689319 885 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8345149 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6249949 5', mRNA sequence.
ACCESSION BQ689319
VERSION BQ689319.1 GI:21814635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM2392 row: d column: 14
High quality sequence stop: 667.
FEATURES Location/Qualifiers
source 1. .885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6249949"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 32.3%; Score 707; DB 5; Length 885;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 807; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       361 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11 CCATCGAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGA 70

Qy       421 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        71 GGTTGATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCC 130

Qy       481 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       131 AGAAAGAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGA 190

Qy       541 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       191 AGACCCTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATG 250

Qy       601 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       251 TCAAGGCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTG 310

Qy       661 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       311 CCGGAAAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCA 370

Qy       721 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       371 CCTTGCACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTG 430

Qy       781 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       431 GCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGA 490
```

```
Qy        841 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 TCCAAGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAGCAGC 550

Qy        901 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 TGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACCTGG 610

Qy        961 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 TGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCATCA 670

Qy       1021 CCCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCAAGACAAGG 1080
              | |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        671 CTCTCGAAGTGGAGCCGAGTGACACCATTGAGAATGTCAAGGCAAAGATCCNAGACAAGG 730

Qy       1081 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 AAGGCATCCCTCCTGACCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATGGTC 790

Qy       1141 GTACCCTGTCTGACTACAACATCCAGAAA 1169
              |||||||||||||||||||||||||||||
Db        791 GTACCCTGTCTGACTACAACATCCAGAAA 819
```

RESULT 260
CR982652
LOCUS CR982652 896 bp mRNA linear EST 24-JUN-2005
DEFINITION CR982652 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016I0812 5', mRNA sequence.
ACCESSION CR982652
VERSION CR982652.1 GI:68220986
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D. and Korn,B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDp9016I0812.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016I0812
contact RZPD (product- support@rzpd.de) for further information.
Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES Location/Qualifiers
source 1. .896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016I0812"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"
ORIGIN

Query Match 32.3%; Score 707; DB 7; Length 896;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 807; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 111

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 171

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 231

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 291

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        292 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 351

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 411

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 471

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 531

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 591

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 651

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 711

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 771

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 845
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 ANCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGAGAACGTCAAAGCAAAGATCCAA 831

Qy        846 GACAAGGAAGGCATTCCTCCTGACCAGCA 874
              |||||||||||||||||||||||||||||
Db        832 GACAAGGAAGGCATTCCTCCTGACCAGCA 860
```

RESULT 261
CR985699
LOCUS CR985699 911 bp mRNA linear EST 28-JUN-2005
DEFINITION CR985699 RZPD no.9017 Homo sapiens cDNA clone RZPDp9017B2231 5',

```
            mRNA sequence.
ACCESSION   CR985699
VERSION     CR985699.1  GI:68279584
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 911)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9017B2231.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9017
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9017 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9017B2231
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .911
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDp9017B2231"
                     /tissue_type="T-Lymphocytes"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="RZPD no.9017"
                     /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
                     NotI; vector:
                     http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
                     ; 1st strand cDNA was prepared from mRNA obtained from
                     human T-Lymphocytes with a NotI - oligo(dT) primer [5'
                     GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to SalI adaptors,
                     digested with NotI and cloned into the NotI and SalI sites
                     of the pQE80LSN_cloned vector"
ORIGIN

  Query Match             32.3%;  Score 707;  DB 7;  Length 911;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  757;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 114

Qy        126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 174

Qy        186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 234

Qy        246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 294

Qy        306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        295 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 354

Qy        366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 414

Qy        426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 474

Qy        486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 534

Qy        546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 594

Qy        606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 654

Qy        666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 714

Qy        726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 774

Qy        786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
              |||||||||||||||||||||||||||||||||||||
Db        775 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 812


RESULT 262
CT002282
LOCUS       CT002282                913 bp    mRNA    linear   EST 28-JUN-2005
DEFINITION  CT002282 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016B2223 5',
            mRNA sequence.
ACCESSION   CT002282
VERSION     CT002282.1  GI:68296165
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
  TITLE     Human T-Lymphocytes library
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; RZPDp9016B2223.
            RZPDLIB; (Human T-Lymphocytes) RZPD LIB No.9016
            http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
            Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            http://www.rzpd.de/cgi- bin/products/cl.cgi?CloneID=RZPDp9016B2223
            contact RZPD (product- support@rzpd.de) for further information.
            Primer name: qe3_4 , Primer sequence: CGGATAACAATTCACACAG.
FEATURES             Location/Qualifiers
     source          1. .913
```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016B2223"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"

ORIGIN

Query Match 32.3%; Score 707; DB 7; Length 913;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        66 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        51 ACAATGCAGATCTTCGTGAAGACTCTGACTGGTAAGACCATCACCCTCGAGGTTGAGCCC 110

Qy       126 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       111 AGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATAAGGAAGGCATCCCTCCTGAC 170

Qy       186 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       171 CAGCAGAGGCTGATCTTTGCTGGAAAACAGCTGGAAGATGGGCGCACCCTGTCTGACTAC 230

Qy       246 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       231 AACATCCAGAAAGAGTCCACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAAATC 290

Qy       306 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTGGAGCCCAGTGACACCATC 365
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       291 TTCGTGAAGACACTCACTGGCAAGACCATCACCCTTGAGGTCGAGCCCAGTGACACCATC 350

Qy       366 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       351 GAGAACGTCAAAGCAAAGATCCAGGACAAGGAAGGCATTCCTCCTGACCAGCAGAGGTTG 410

Qy       426 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       411 ATCTTTGCCGGAAAGCAGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAA 470

Qy       486 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       471 GAGTCTACCCTGCACCTGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACC 530

Qy       546 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       531 CTGACTGGTAAGACCATCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAG 590

Qy       606 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       591 GCAAAGATCCAAGATAAGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGA 650

Qy       666 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       651 AAACAGCTGGAAGATGGTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTG 710

Qy       726 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       711 CACCTGGTACTCCGTCTCAGAGGTGGGATGCAAATCTTCGTGAAGACACTCACTGGCAAG 770

Qy       786 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 823
             |||||||||||||||||||||||||||||||||||||
Db       771 ACCATCACCCTTGAGGTCGAGCCCAGTGACACTATCGA 808
```

RESULT 263
BQ876710
LOCUS BQ876710 916 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8584440 Lupski_sympathetic_trunk Homo sapiens cDNA clone IMAGE:6192931 5', mRNA sequence.
ACCESSION BQ876710
VERSION BQ876710.1 GI:22268718
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 916)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM13595 row: 1 column: 20
High quality sequence stop: 624.

FEATURES Location/Qualifiers
source 1. .916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6192931"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sympathetic_trunk"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

ORIGIN

Query Match 32.3%; Score 707; DB 5; Length 916;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       442 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AGCTGGAAGATGGGCGCACCCTGTCTGACTACAACATCCAGAAAGAGTCTACCCTGCACC 60

Qy       502 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TGGTGCTCCGTCTCAGAGGTGGGATGCAGATCTTCGTGAAGACCCTGACTGGTAAGACCA 120

Qy       562 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TCACCCTCGAGGTGGAGCCCAGTGACACCATCGAGAATGTCAAGGCAAAGATCCAAGATA 180

Qy       622 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AGGAAGGCATTCCTCCTGATCAGCAGAGGTTGATCTTTGCCGGAAAACAGCTGGAAGATG 240
```

```
Qy      682 GTCGTACCCTGTCTGACTACAACATCCAGAAAGAGTCCACCTTGCACCTGGTACTCCGTC 741
            ||||
```

THIS PAGE BLANK (USPTO)